Practical Methods for Assessing Image Quality in Diagnostic Radiology

Nicholas Marshall

Department of Radiology, University Hospitals Leuven, Herestraat 49, 3000 Leuven, Belgium

general x-ray

conebeam dental

intra-oral

dynamic imaging (fluoroscopy)

Computed Tomography

mammography

Diagnostic x-ray systems

Diagnostic X-ray Imaging Systems

- Three main classes of system
- Projection imaging
  - General radiography, intraoral dental, mammography
- Dynamic projection imaging
  - Fluoroscopy, cardio-angiography, digital subtraction angiography
- Volumetric imaging
  - CT, conebeam CT, digital breast tomosynthesis, rotational angiography
- How do we evaluate the quality of the images produced by these systems?

Image Quality – What is it?

- "A scientific or medical image is always produced for some specific purpose or task, and the only meaningful measure of its quality is how well it fulfills that purpose" - Barrett (1990)

Receiver Operating Characteristic (ROC) curve and variants ⇔ **Visual Grading Analysis (VGA)** ⇔ **Alternative Forced Choice (AFC)** ⇔ **Contrast-detail curve Line pair TO** ⇔ **Quantitative Measurements MTF [u] NNPS [u] DQE [u]**

CLINICAL (observers) ⟺ **TECHNICAL (numbers)**

Why are we trying to measure image quality of our system?

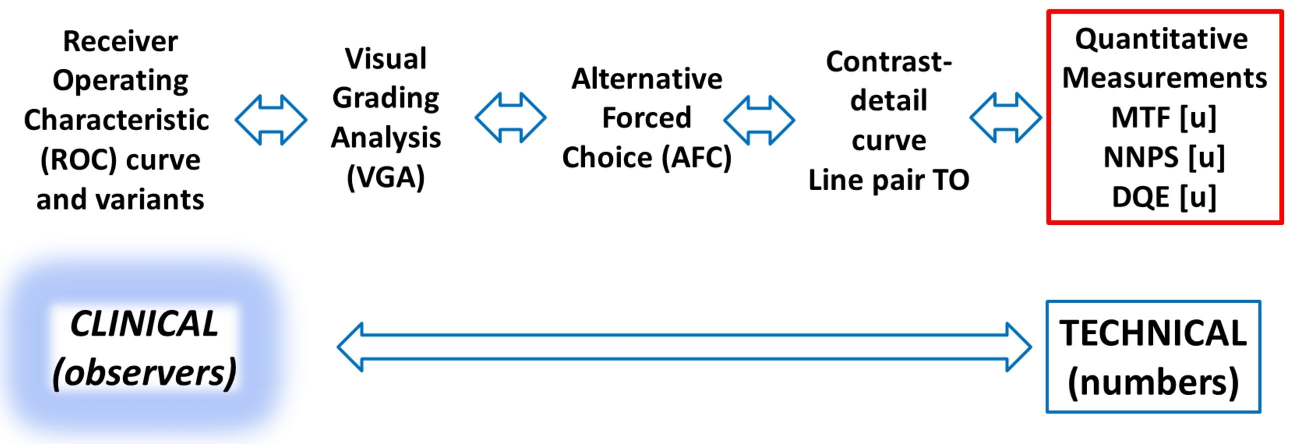
- We cannot perform a clinical study everytime we want to assess image quality…in fact not required in many situations
- The motivation for the measurement largely defines the type of image quality assessment made
- Quality Control
  - “Has it changed?” -> system performance changed?
- Optimization
  - “Is it optimized?” -> system sitting at optimal point?
- Performance
  - “Is it any good?”-> system suitable for my clinical tasks?

Quality Control

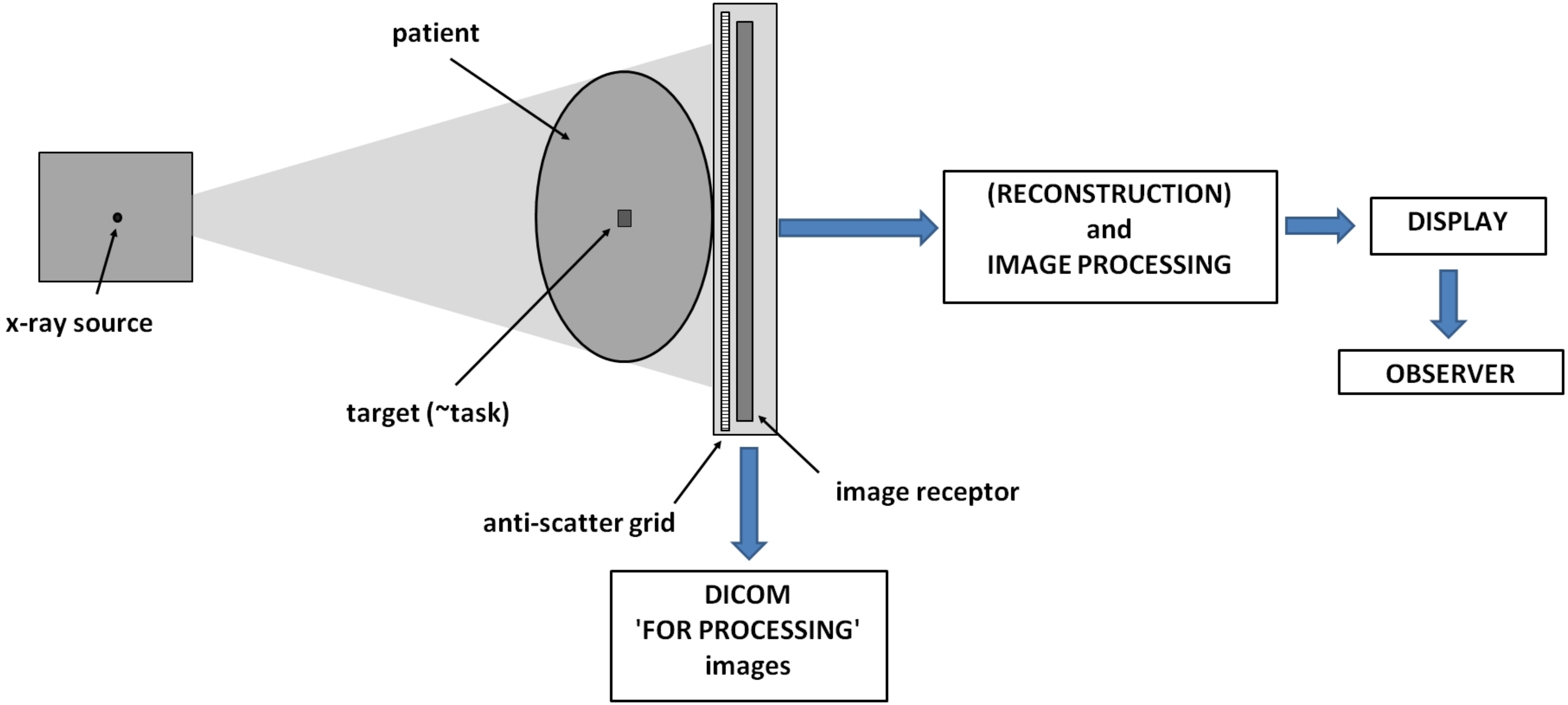
- “Has it changed?” - > Classic QC
- Pick parameters to monitor over time
- The QC cycle begins:
  - QC Measurement
  - Compare -> baselines
  - Correct or replace component
- Best way to do this depends on:
  - Modality
  - Access to images and the image types available
  - Test objects available

Quality Control

- Rather than a direct measure of image quality, we asssess parameters that are synonymous with good image quality
  - Sharpness, Contrast, Noise
  - Dynamic range, Uniformity
  - Lag and ghosting
- These parameters can be readily assessed using test objects and/or quantitative measurements



Imaging system



Quality Control - Technical characterization

- Characterize according to International Electrotechnical Commission (IEC) document IEC 62220-1
  - Detector response curve
  - Presampling Modulation Transfer Function (MTF) to quantify detector sharpness
  - Normalized Noise Power Spectrum (NNPS) to quantify noise
  - Detective Quantum Efficiency (DQE) to quantify detector signal-to-noise ratio (SNR) transfer efficiency

$$DQE(u) = \frac{MTF^2(u)}{K_a\, Q_{in}\, NNPS(u)}$$

INTERNATIONAL STANDARD

IEC 62220-1

First edition
2003-10

**Medical electrical equipment –
Characteristics of digital X-ray imaging devices –
Part 1: Determination of the detective quantum efficiency**

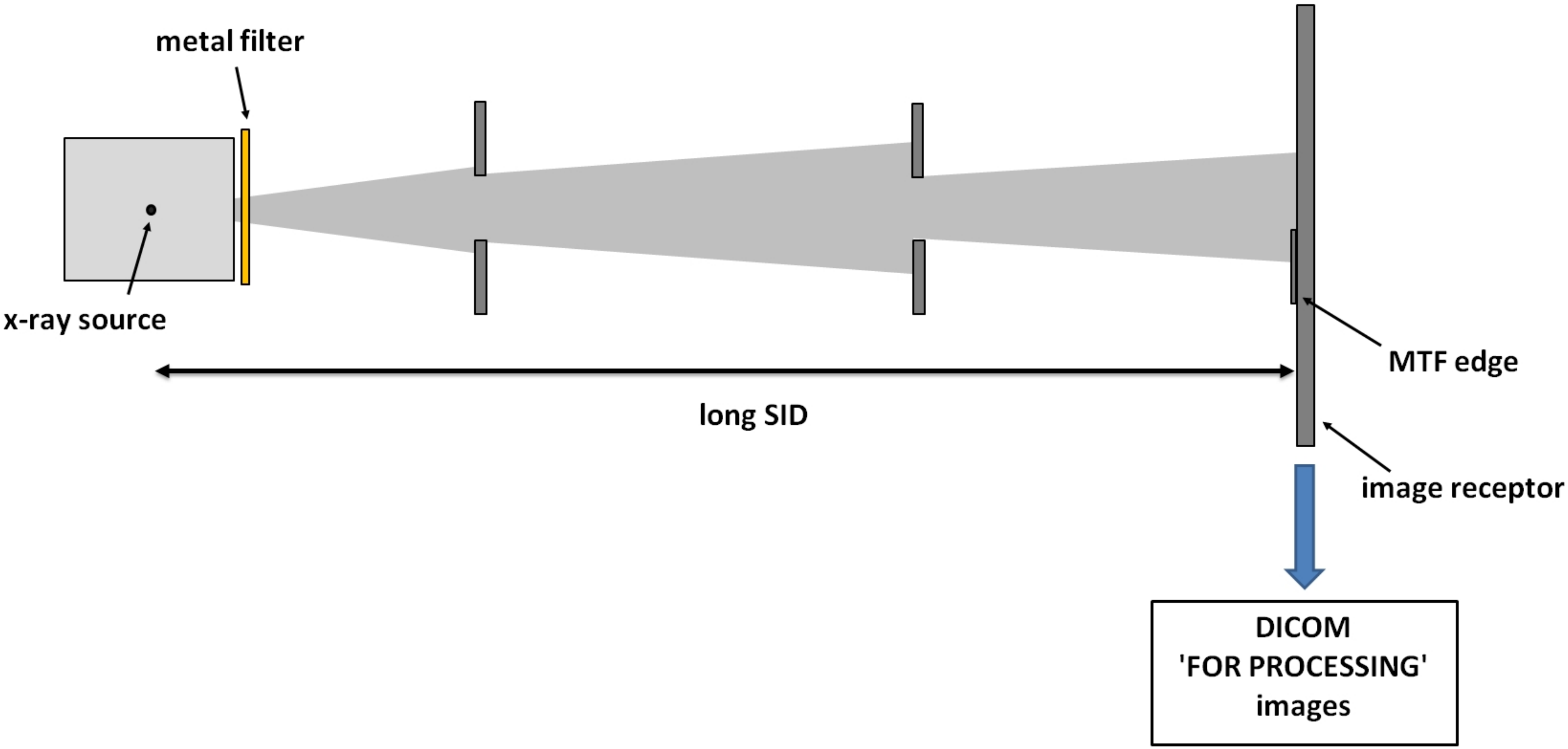
*Appareils électromédicaux –
Caractéristiques des appareils d'imagerie à rayonnement X –
Partie 1: Détermination de l'efficacité quantique de détection*

IEC

Reference number
IEC 62220-1:2003(E)

Quality Control – quantitative measurements

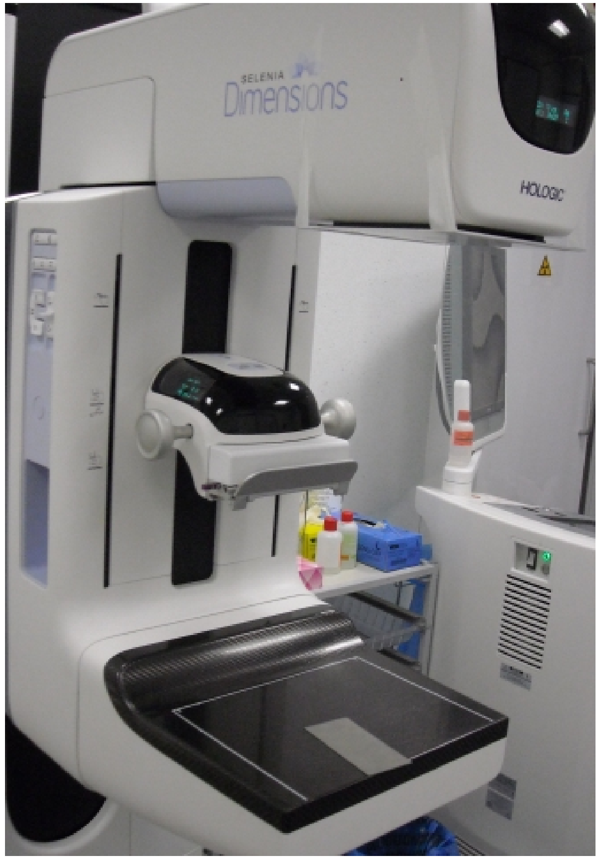
- Explicitly a detector test



Output, detector response and flood images

X-ray output

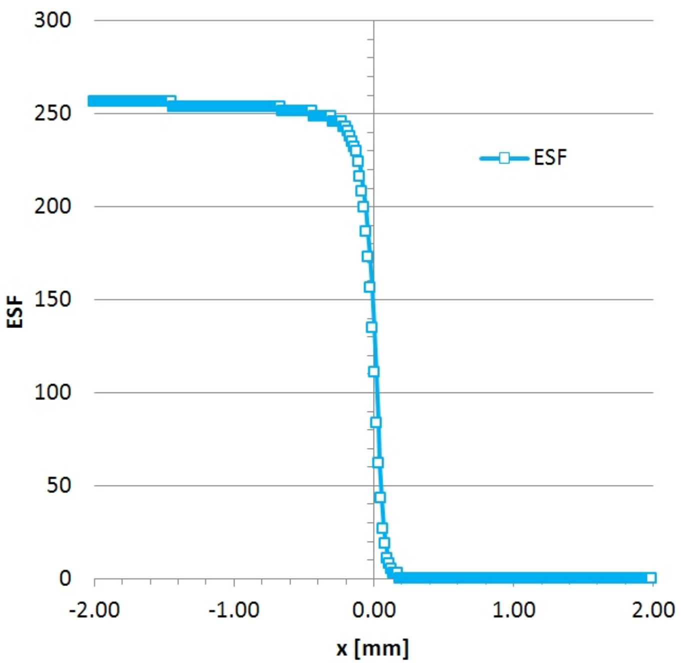
Homogenous images

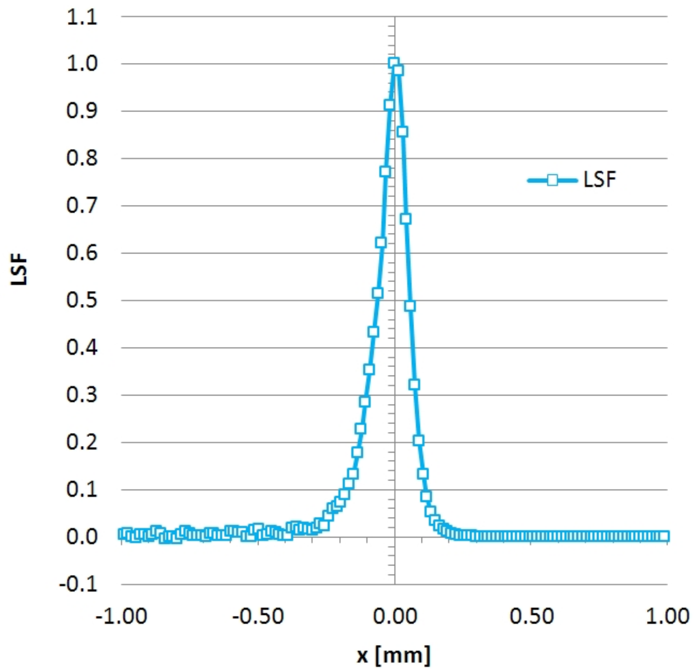


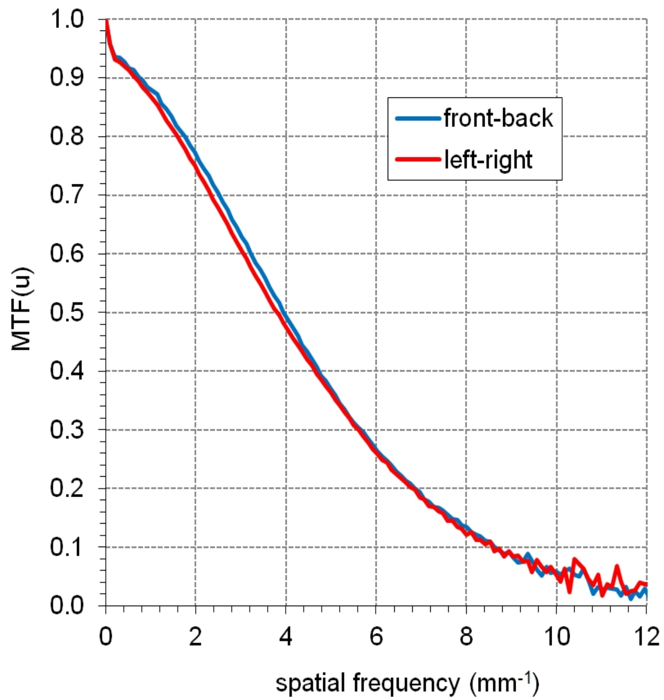
MTF images

MTF using the edge method

- Measured in 2 directions across the detector





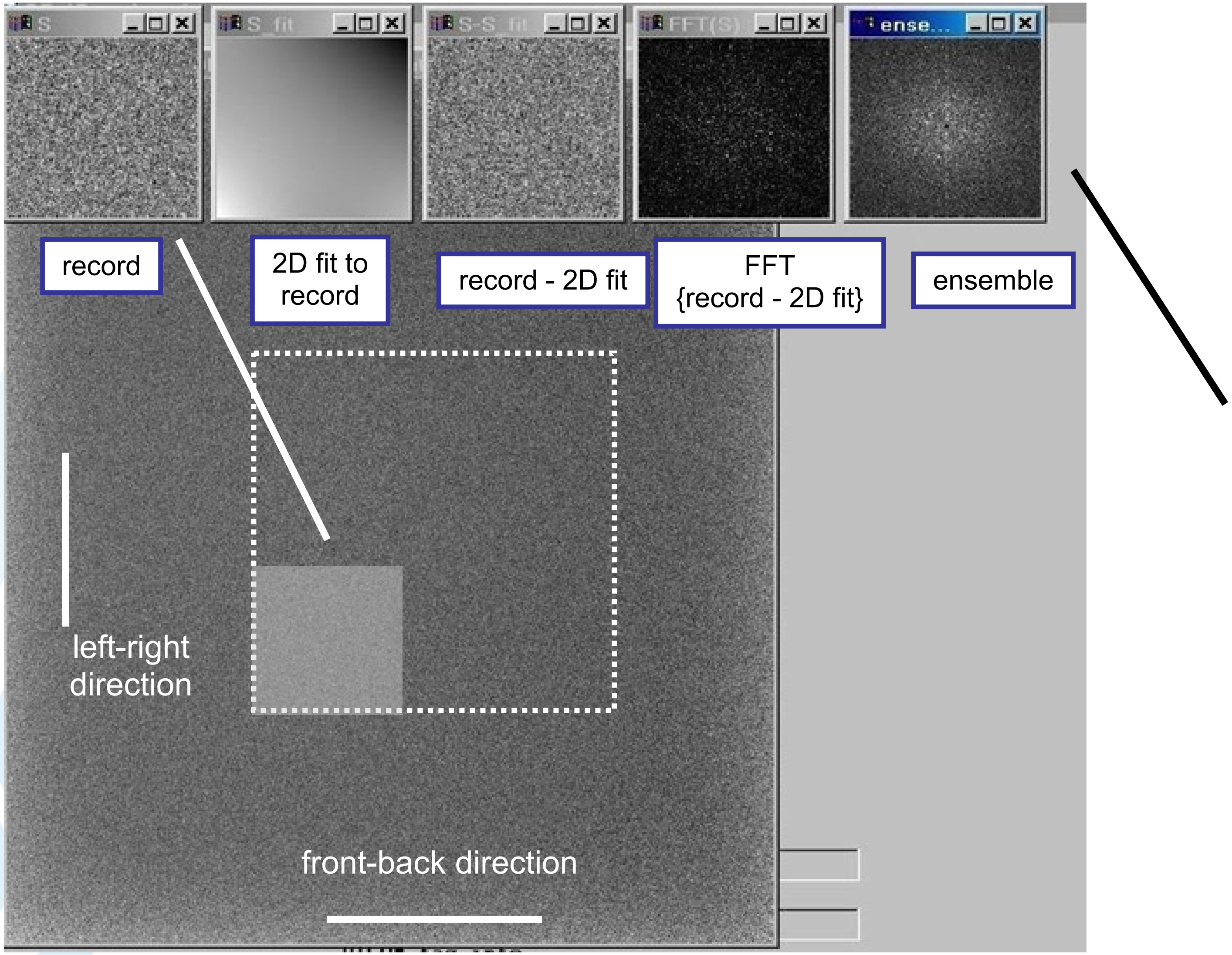


Edge Spread Function — Line Spread Function — Pre-sampling MTF

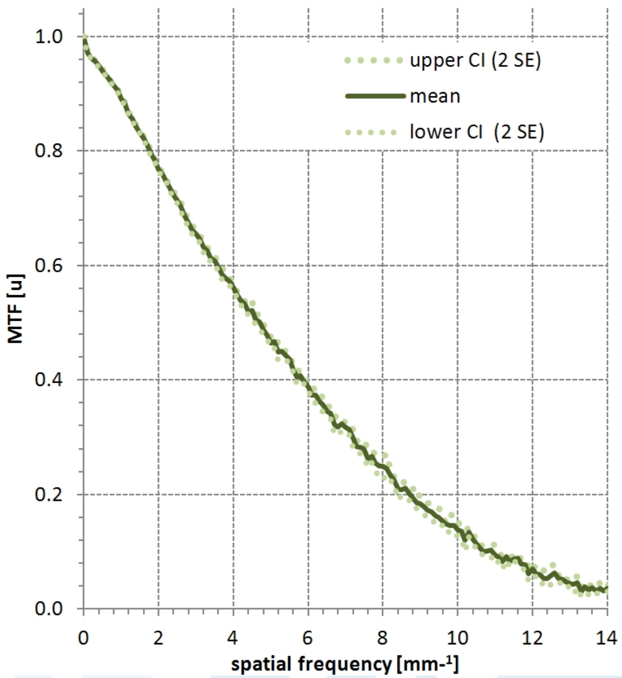
differentiate *FFT*

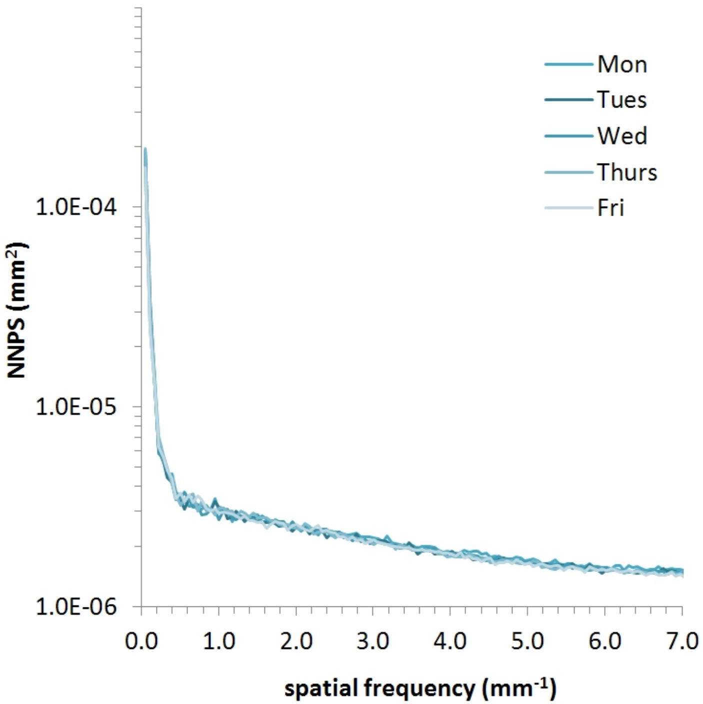
Samei E, Flynn M J and Reimann D A 1998 A method for measuring the presampled MTF of digital radiographic systems using an edge test device *Med. Phys.* **25** 102–13

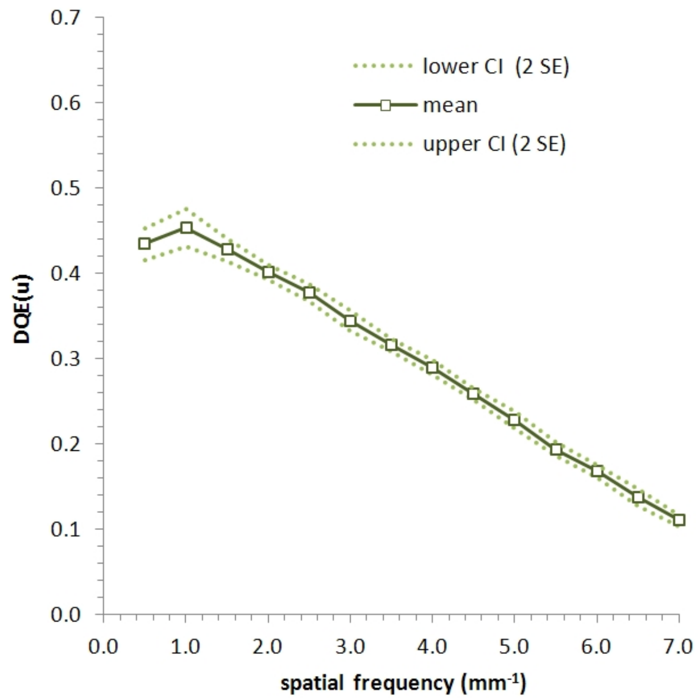
NPS – calculate from homogeneous images



QC - Repeatability of quantitative measures







- MTF cov at 1 mm^{-1} 0.2%
- MTF cov at 5 mm^{-1} 2.3%
- MTF cov at 10 mm^{-1} 9.0%
- MTF mean cov up to Nyquist = 1.6%

- NNPS cov < 1 mm^{-1} ~5%
- NNPS mean cov = 3.2%

- DQE mean cov = 4.4%

<u>Measured on 5 consecutive days for mammo system</u>

QC – Long term stability of quantitative measures

- Reproducibility over 17 months for a given system

cov for 50% MTF ~3%

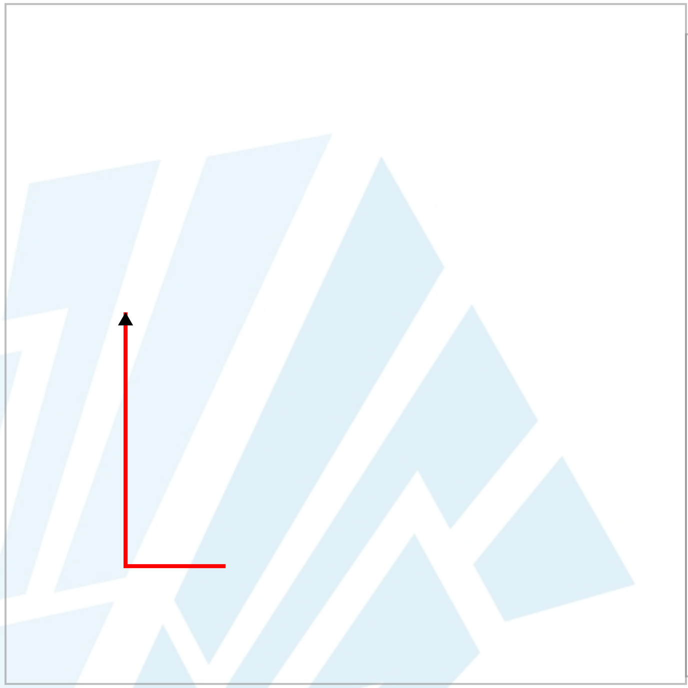
cov for NNPS at 2 mm^{-1} ~4%

cov for DQE at 1 mm^{-1} ~7%

Marshall NW Early experience in the use of quantitative image quality measurements for the quality assurance of full field digital mammography x-ray systems *Phys. Med. Biol.* **52** (2007) 5545–5568

QC – general applicability of measures

- Application to 6 identical CsI based DR detectors



cov for 50% MTF was 3.9%

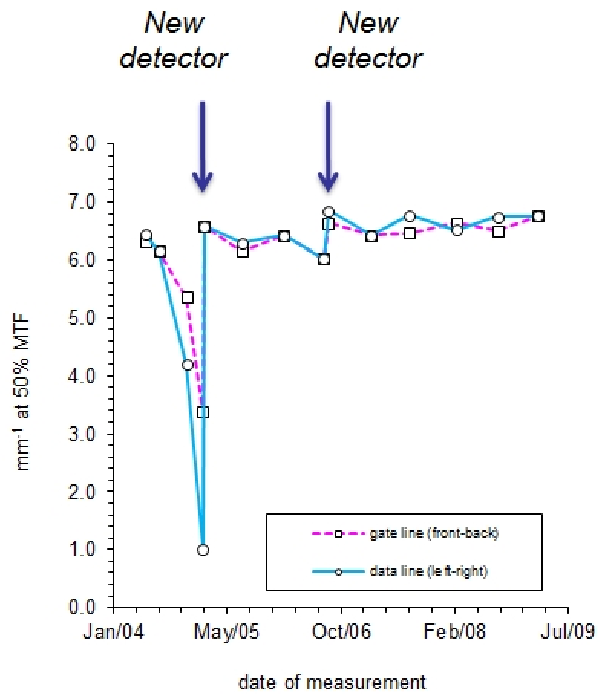
cov for NNPS was 6.5%

cov for DQE ~2 - 10%

Marshall NW, Mackenzie A and Honey I Quality control measurements for digital x-ray detectors *Phys. Med. Biol.* **56** (2011) 979-999

Quantitative measurements: sensitivity

- Sensitive to changes in detector sharpness: a-Se crystallization



front-back direction

left-right direction

Marshall NW 2006 Retrospective analysis of a detector fault for a full field digital mammography system *Phys. Med. Biol.*

QC – track changes

- Noise power spectrum has an MTF dependency
- Sensitive to detector blurring

front-back direction

left-right direction

faulty detector

healthy detector

NNPS at ~100 μGy

QC – trouble shooting using quantitative measures

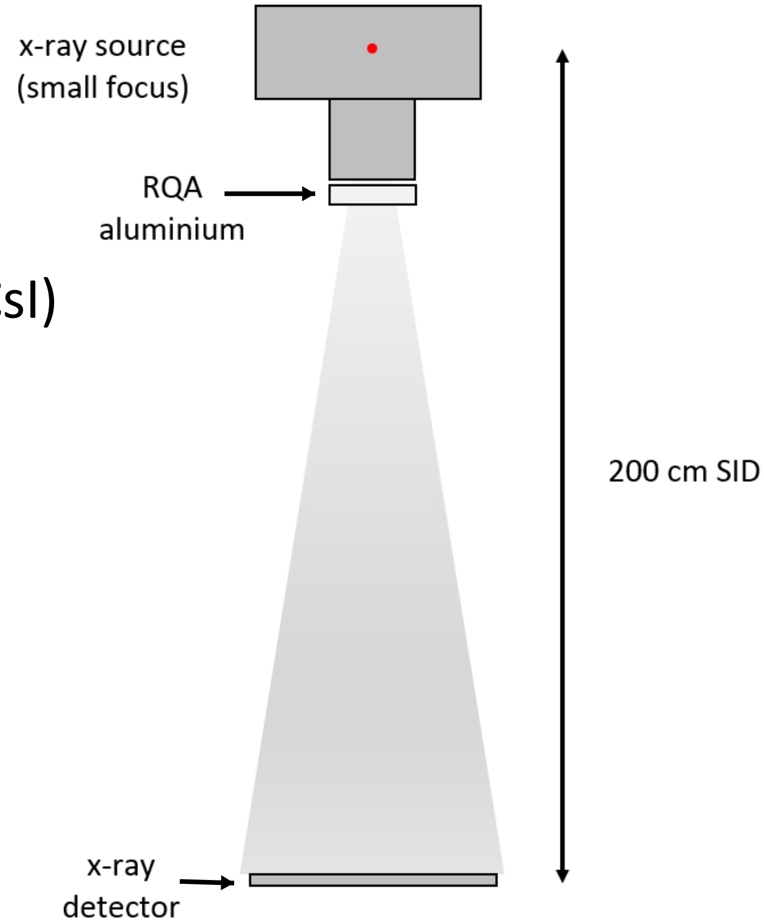
- "We don't like the images produced by the mobile system with Carestream wireless DR detector on the neonate ward"
- "How good are the new needle CR detectors"?
- Where to start? ← **kV accuracy, HVL and output ok…**
  - X-ray source
  - X-ray detector
  - Image processing
- Detector?

Receiver Operating Characteristic (ROC) curve and variants ⇔ **Visual Grading Analysis (VGA)** ⇔ **Alternative Forced Choice (AFC)** ⇔ **Contrast-detail curve Line pair TO** ⇔ **Quantitative Measurements MTF [u] NNPS [u] DQE [u]**

CLINICAL (observers) ⇔ **TECHNICAL (numbers)**

QC – trouble shooting using quantitative measures

- Technical characterization of four detectors
  - Agfa CR powder phosphor MD4.0 (BaFBr)
  - Agfa CR needle phosphor HD5.0 (CsBr)
  - Agfa DR flat panel detector DX-D35C (CsI)
  - Carestream DR flat panel detector DRX 2530C (CsI)

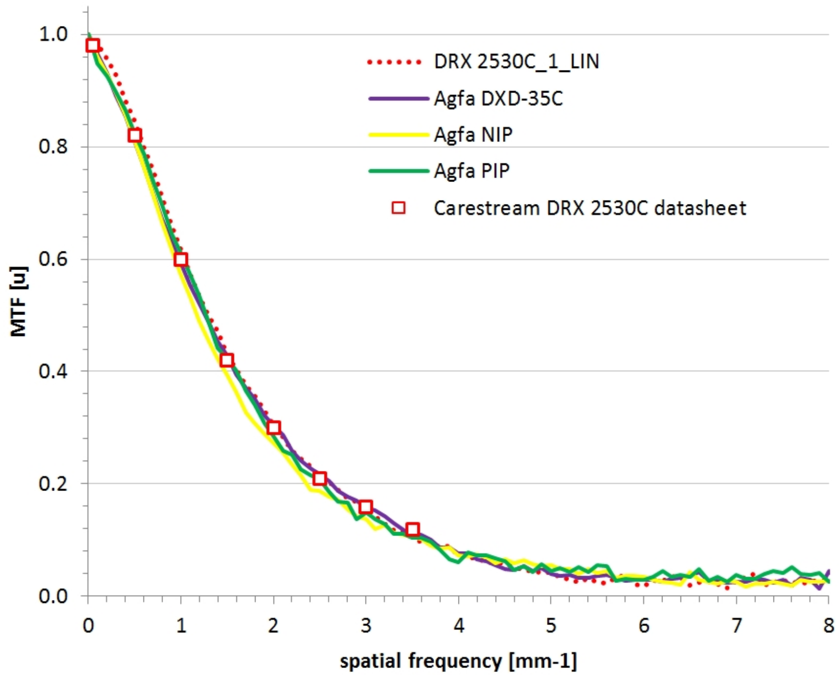


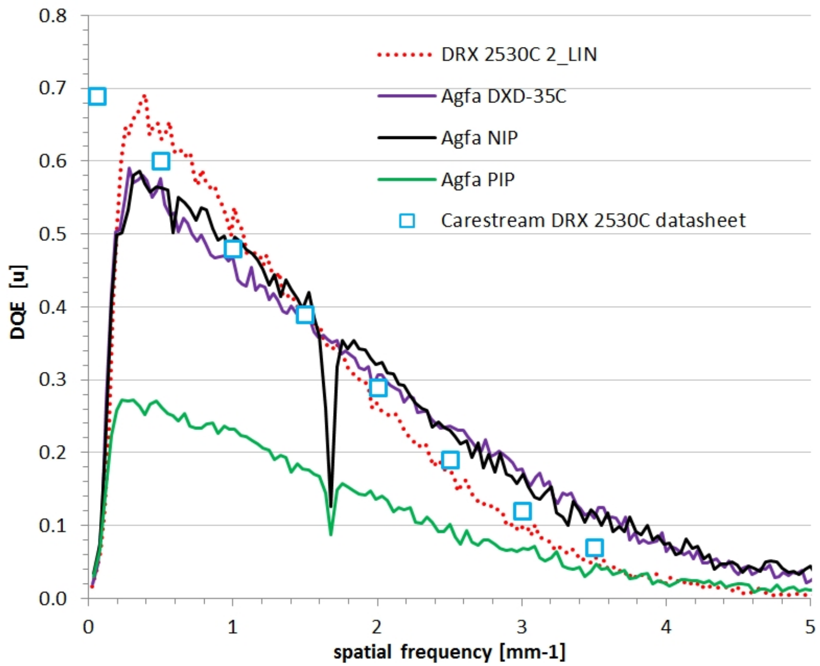
QC – trouble shooting using quantitative measures

- DICOM ‘FOR PROCESSING’ images
  - Agfa FLATFIELD 400
- Response Function
  - Homogeneous images as fn detector air kerma (K)
- MTF edge
  - 50x50 mm Tantalum (Ta) edge, angled ~3°
  - 4 images, 90° rotation between images
  - Imaged at 8.0 µGy
- NNPS from the homogeneous images
  - 0.7 µGy, 2.5 µGy, 8.0 µGy target K
  - 3 images at each K level

QC – trouble shooting using quantitative measures

- Carestream detector has good/best efficiency – not a detector problem
- Needle CR has clearly superior DQE compared to Powder CR
- Needle CR and DR detectors have similar efficiency





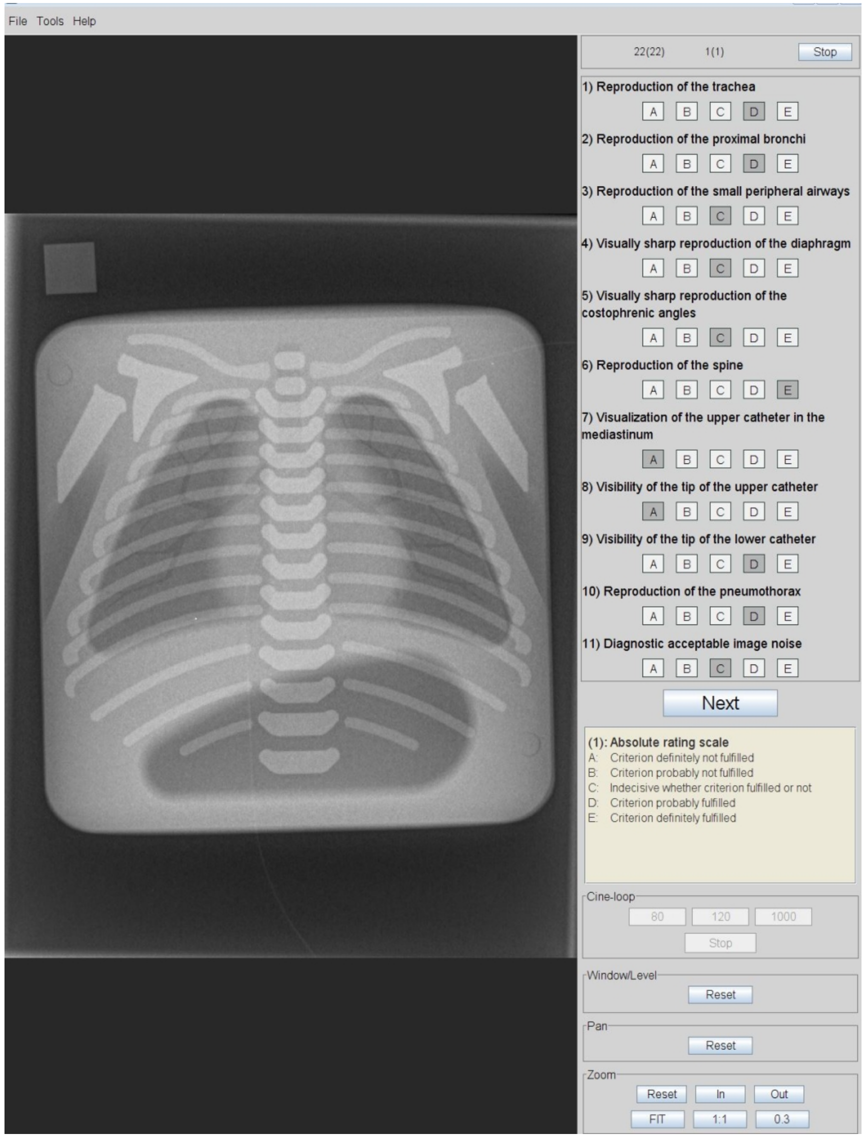
RQA3 beam quality [52 kV, 10 mm Al]

QC – trouble shooting using quantitative measures

- What next? Image processing….
- Carestream changed neonate chest processing
- Default was USA preset [~smooth and with high contrast]
- Changed to match UZ Leuven taste [sharp with more noise]
- Visual Grading Analysis (VGA) using a Gammex 610 phantom
  - ~1.5 kg neonate
  - 2 catheters added to the phantom ask a typical task
  - Seven dose levels around clinical default of 63 kV; 0.8 mAs; 100 cm SID, no antiscatter grid
  - Agfa MUSICA^2 and Carestream neonate processing…

QC – trouble shooting using quantitative measures

- ViewDex software used…
- VGA results
  - Carestream performed just ahead of Agfa DR and CR
  - Agfa DR and needle CR similar
  - Agfa powder CR poorest
- Results support the DQE results



Receiver Operating Characteristic (ROC) curve and variants ⇔ **Visual Grading Analysis (VGA)** ⇔ Alternative Forced Choice (AFC) ⇔ Contrast-detail curve Line pair TO ⇔ Quantitative Measurements MTF [u] NNPS [u] DQE [u]

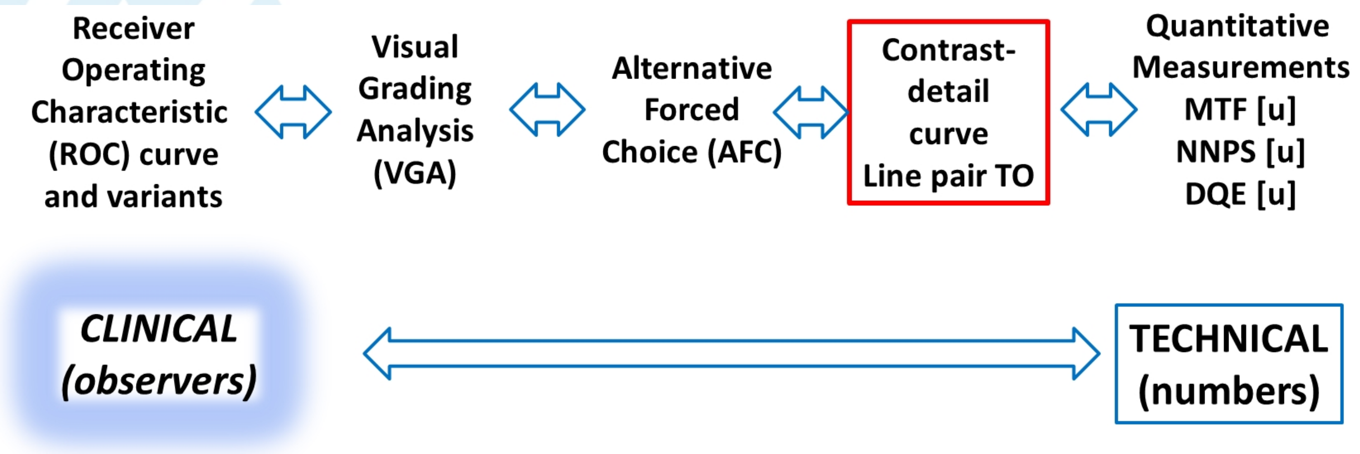
CLINICAL (observers) ⟷ TECHNICAL (numbers)

Quantitative measurements - summary

- Quantitative measures are an excellent means of assessing:
  - Absolute detector performance, changes in performance
  - Sensitive and reproducible
- DQE absolute measure of detector efficiency under ~'bench' conditions
- These measures are 'task neutral' – they tell us about the 'health' of the detector, the potential for image quality, but not image quality per se…
- We have to introduce a task ….. e.g. using a test object

Test Objects

- Test objects
  - Classic methods, carried over from analogue system testing
- Sharpness -> line pair test object
  - Still in use (even for CT) but of limited use for digital detectors
- Basic tasks
  - Objects of interest on a uniform background
  - Threshold contrast-detail (c-d) curve



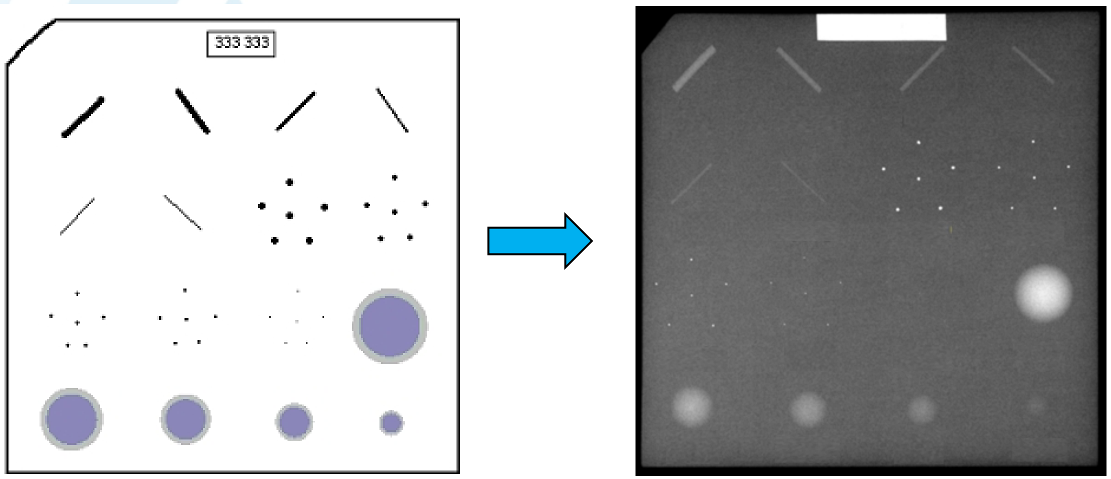
Test Object IQ

- Widely used by physicists for all modalities
- E.g. mammography testing in North America
  - American College of Radiology (ACR) Accreditation phantom
  - Basic targets: fibers, specks, masses
  - Human observer scoring

fibers

specks

masses



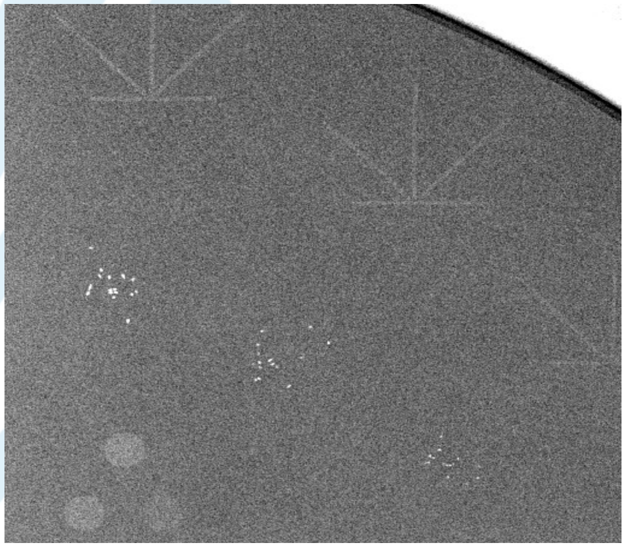
At a minimum,

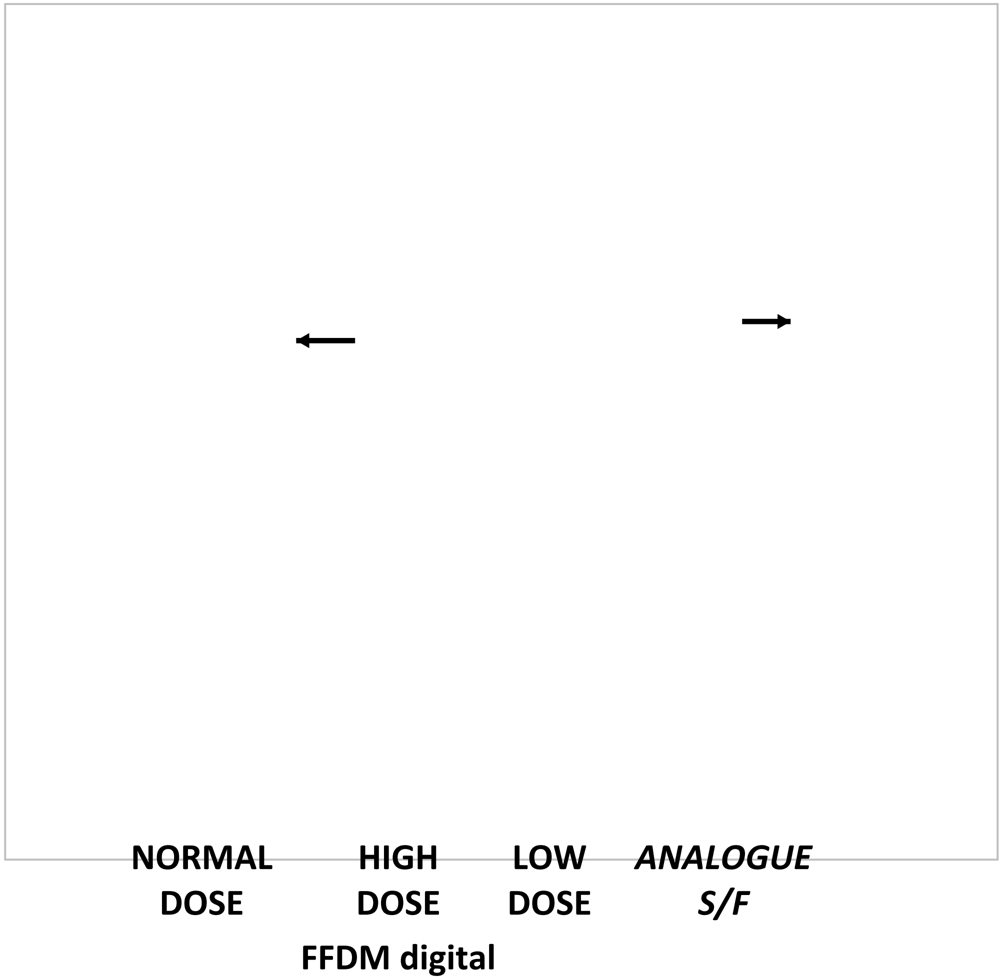
four largest fibers,
three largest speck groups,
three largest masses

must be visible

Test Object IQ

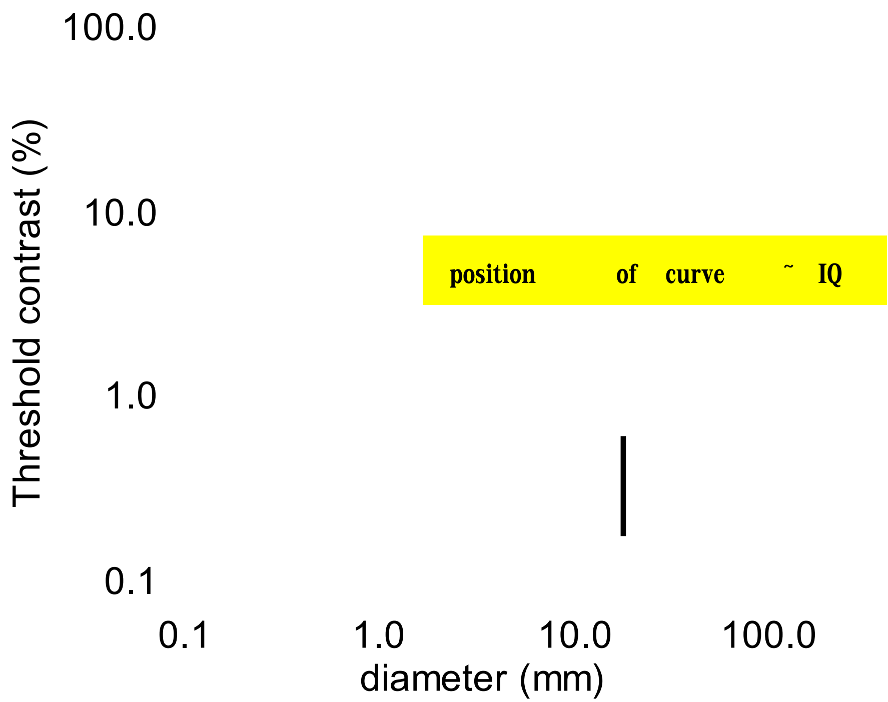
- Leeds TORMAM mammography test object
  - used in Europe
  - Again using basic targets
  - Subjective but follows performance (changes in noise and contrast)





Test Object IQ - Threshold contrast-detail

- Threshold contrast detail (c-d)
- Used for radiography and fluoroscopy system evaluation
- Leeds TO10 and TO20



Test object made from Al, Cu or Pb discs — x-ray image — contrast–detail graph

Hay et al *Brit J Radiol* (1985), Cowen et al *Brit J Radiol* (1987)

Test Object IQ - Threshold contrast-detail

- processing c-d results
  - Calculate contrasts for beam quality used (look up table from Leeds manual)
  - Contrast is 'removed' from the evaluation, this is not a full system test but a detector test

70kV, 1.0 mm Cu

Detail	A	B	C	D	E	F	G	H	J	K	L	M
Size mm	**11.1**	**7.9**	**5.6**	**4**	**2.8**	**2**	**1.4**	**1**	**0.7**	**0.5**	**0.35**	**0.25**
Count 12	6.74	6.74	6.74	8.76	8.76	8.76	23.80	23.80	23.80	95.40	95.40	95.40
11	4.50	4.50	4.50	6.74	6.74	6.74	16.70	16.70	16.70	72.60	72.60	72.60
10	3.22	3.22	3.22	4.50	4.50	4.50	12.80	12.80	12.80	49.60	49.60	49.60
9	2.15	2.15	2.15	3.22	3.22	3.22	8.76	8.76	8.76	36.00	36.00	36.00
8	1.55	1.55	1.55	2.15	2.15	2.15	6.74	6.74	6.74	23.80	23.80	23.80
7	1.08	1.08	1.08	1.55	1.55	1.55	4.50	4.50	4.50	16.70	16.70	16.70
6	0.86	0.86	0.86	1.08	1.08	1.08	3.22	3.22	3.22	12.80	12.80	12.80
5	0.65	0.65	0.65	0.86	0.86	0.86	2.15	2.15	2.15	8.76	8.76	8.76
4	0.43	0.43	0.43	0.65	0.65	0.65	1.55	1.55	1.55	6.74	6.74	6.74
3	0.30	0.30	0.30	0.43	0.43	0.43	1.08	1.08	1.08	4.50	4.50	4.50
2	0.21	0.21	0.21	0.30	0.30	0.30	0.86	0.86	0.86	3.22	3.22	3.22
1	0.16	0.16	0.16	0.21	0.21	0.21	0.65	0.65	0.65	2.15	2.15	2.15

Test Object IQ - Threshold contrast-detail

- Compare against reference curves
  - Evans DS et al 2004 Threshold contrast detail detectability curves for fluoroscopy and digital acquisition using modern image intensifier systems *Brit J Radiol* **77**, 751-758
  - use the appropriate field size
  - need to compare against ref curve at same input dose rate
  - **if the system is quantum noise limited** then scale reference contrast by:

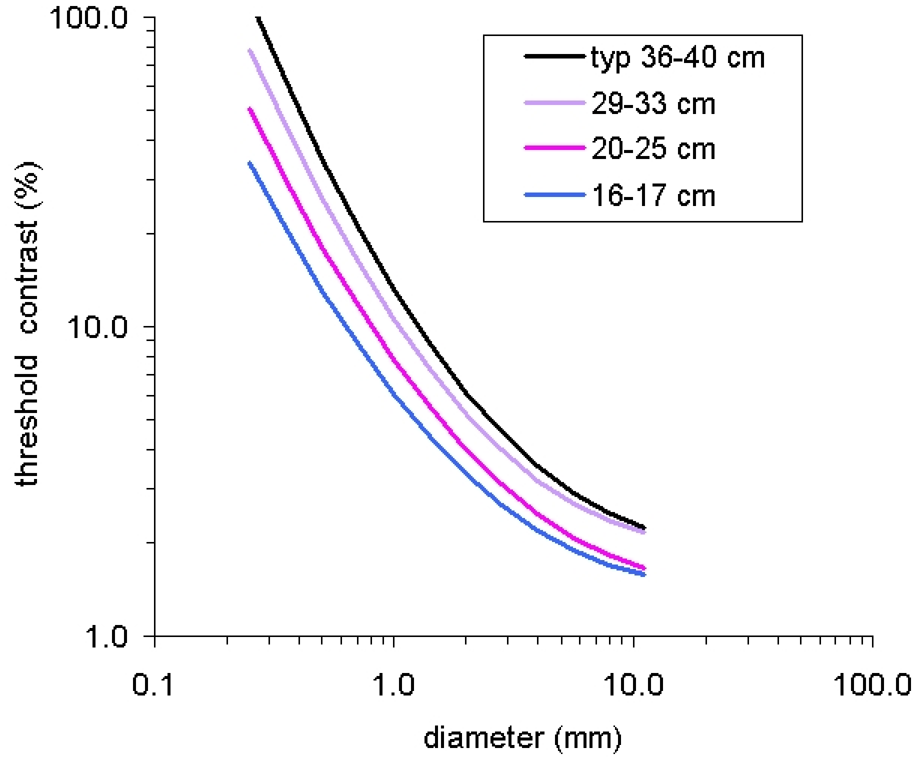
$$C_{ref'} = C_{ref} \times \left(\frac{K_{ref\_curve}}{K_{measurement}} \right)^{0.5}$$

C% reference
scaled to air kerma used for
measurements

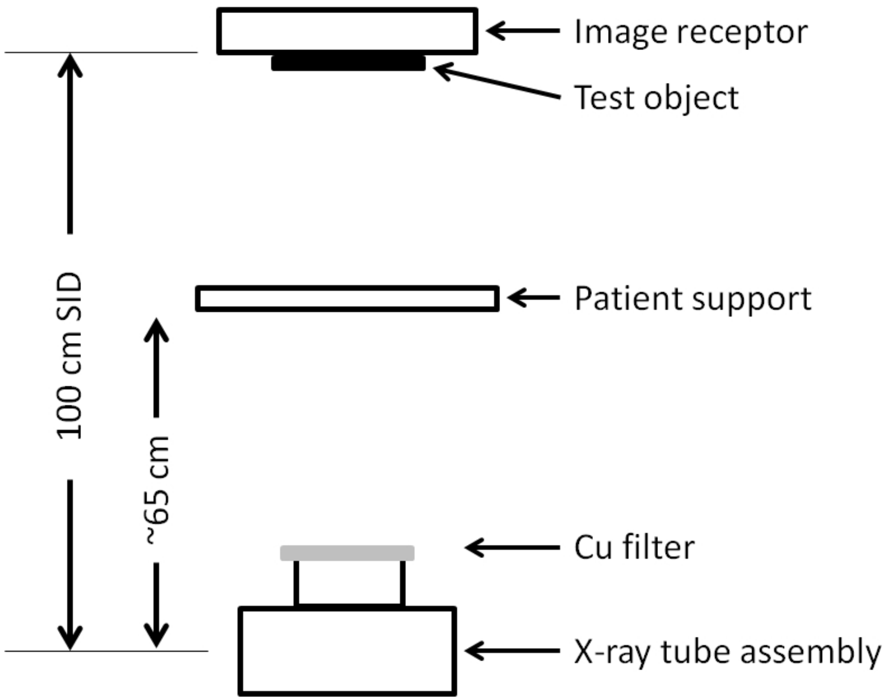
Test Object IQ - Threshold contrast-detail

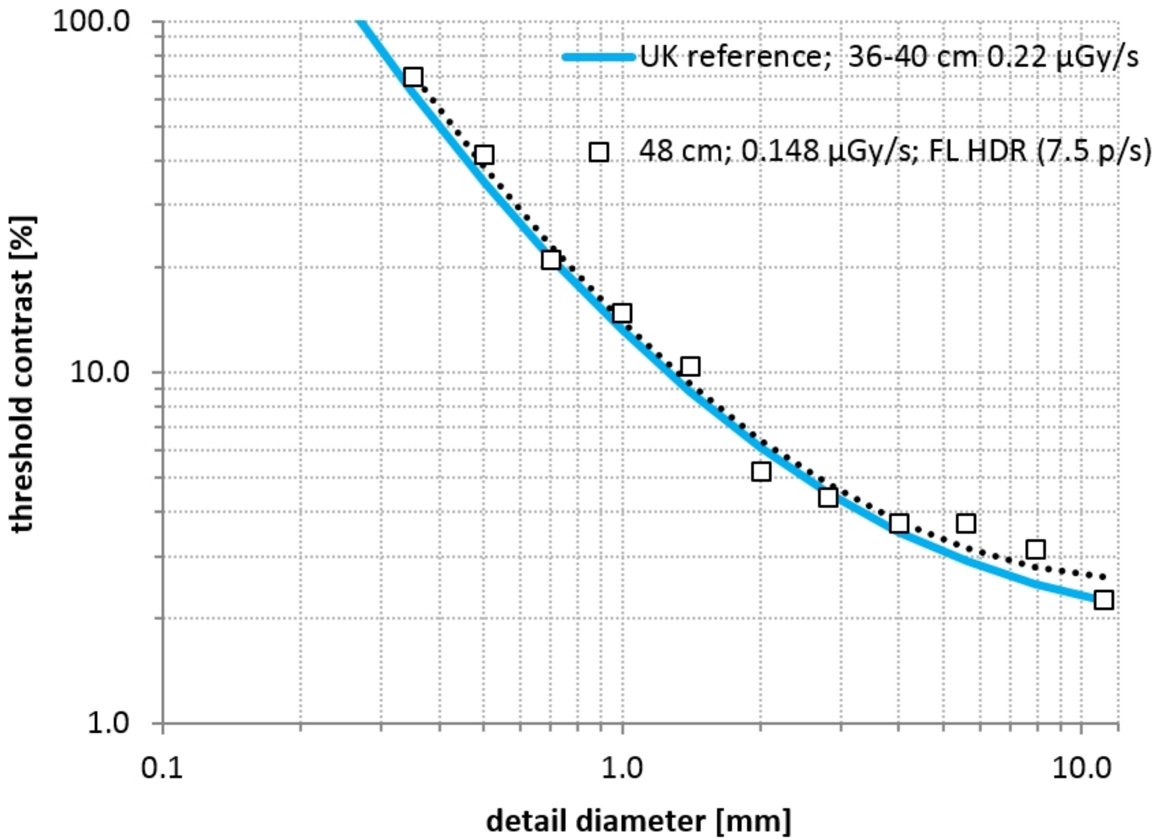
- Evans et al (2004) reference curves

	Threshold Contrast from Evans et al (2004)			
II field size (cm):	*16-17*	*20-25*	*29-33*	*36-40*
II input dose rate (μGy/s):	0.70	0.48	0.27	0.22
[mm]	*[C%]*	*[C%]*	*[C%]*	*[C%]*
11.10	1.59	1.65	2.17	2.24
7.90	1.70	1.82	2.36	2.50
5.60	1.89	2.08	2.67	2.90
4.00	2.18	2.46	3.17	3.52
2.80	2.65	3.09	4.00	4.56
2.00	3.33	4.01	5.23	6.13
1.40	4.43	5.52	7.31	8.86
1.00	6.05	7.79	10.5	13.2
0.70	8.81	11.8	16.4	21.2
0.50	13.1	18.1	26.0	34.9
0.35	20.8	30.0	44.8	62.6
0.25	33.7	50.4	78.5	114.1



Test Object IQ - Threshold contrast-detail





- Siemens Artis Zee flat panel based cardioangiography system
- Detector QC test – detector is compared against typical performance

Performance testing and QC in fluoroscopy (cardiac)

- One of the problems – too many parameters…
- Are our tests sensitive to these parameters?

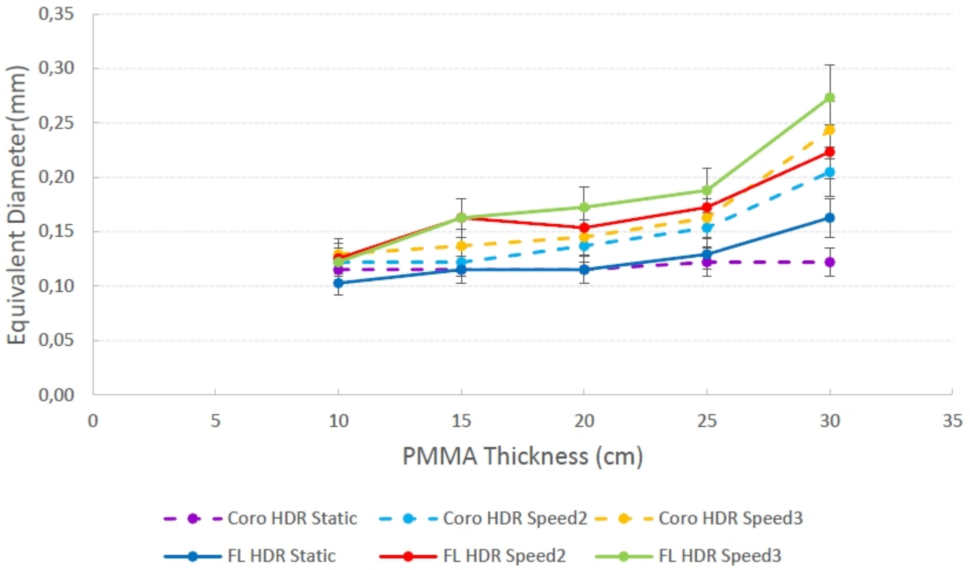
Acquisition	Coro HDR Care	Coro HDR Care minus	LV Car	Clearstent Live LD	Clearstent Live
kV	85 kV	81 kV	81 kV	85 kV	81 kV
Pulse Width	5.0 ms	4.0 ms	5.0 ms	5.0 ms	4.0 ms
kV ms	96 kV	96 kV	Off	102 kV	96 kV
Dose	0.140 μGy/fr	0.080 μGy/fr	0.080 μGy/fr	0.065 μGy/fr	0.170 μGy/fr
kV Dose	102 kV	Off	Off	125 kV	125 kV
Min Cu-filter	0.1 mm	0.1 mm	0.1 mm	0.1 mm	0.0 mm
Max Cu-filter	0.2 mm	0.3 mm	0.2 mm	0.2 mm	0.0 mm
I-Noise Reduction	Normal	Normal	Normal	Normal	Normal
I-Noise Reduction Details	Auto 1	Auto 1	Auto 4	Auto 1	Auto 4
Edge Enhancement NAT	25 %	25%	25 %	20 %	10 %
EE-Kernal	7	5	9	3	3
DDO	40 %	60 %	25 %	60 %	60 %
DDO-Kernal	137	137	201	65	65
K-factor	Auto 5	Auto 5	Off	Auto 1	Off
EVE	Auto 3	Auto 3	Auto 2	Off	Off
Frame Rate	10 fr/s	10 fr/s	30 fr/s	15 fr/s	15 fr/s

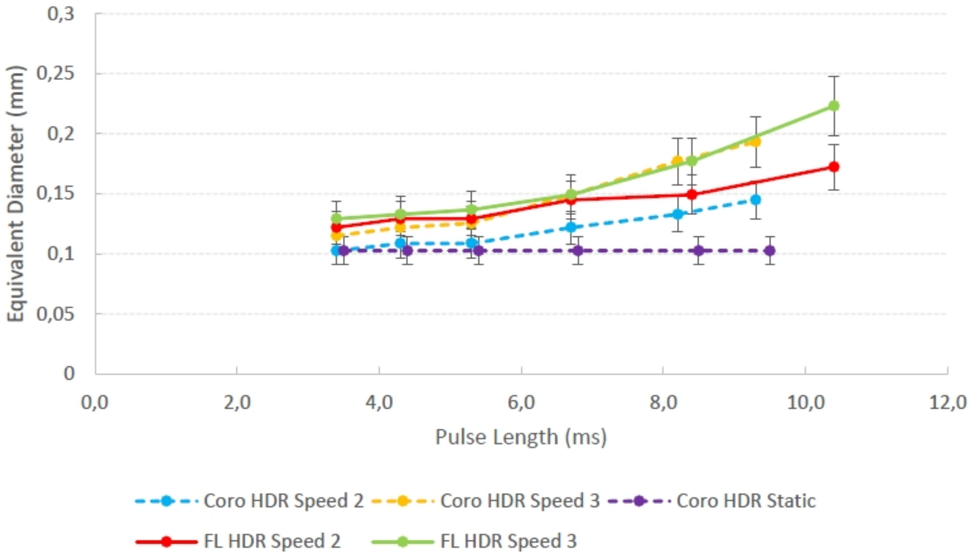
Test Object IQ - in fluoroscopy/cardiac

- Moving lego cart with variable speed setting
  - Cardiac artery velocity varies between 40 mm/s and 90 mm/s
  - Cart velocity at centre of motion = 70 mm/s and 90 mm/s
- Place TO [TO20 c-d or line pair] on cart and score

Performance testing and QC in fluoroscopy (cardiac)

- Convert line pair result to ~diameter that is equivalent to unsharpness



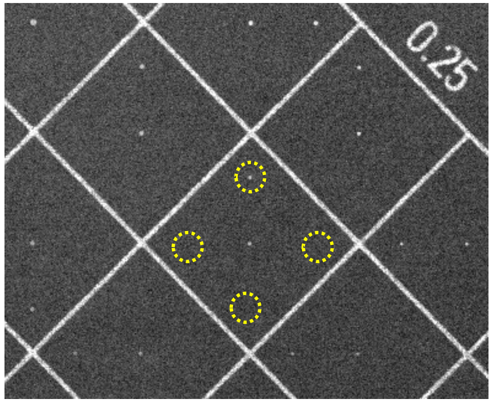


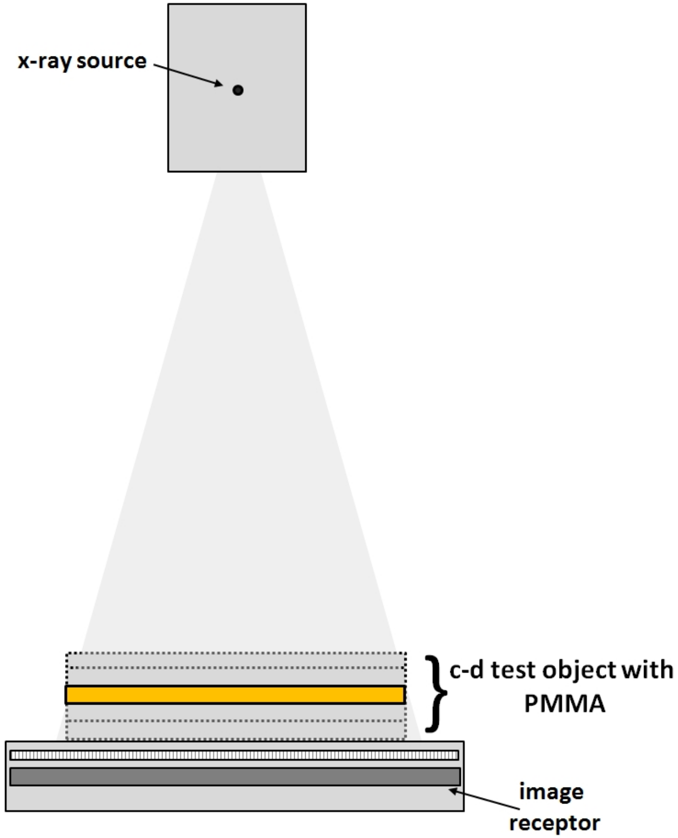
Clinical settings,
vary PMMA thickness

20 cm PMMA,
vary requested pulse length

Test Object IQ - Threshold contrast-detail

- Mammography IQ tests using CDMAM (EUREF)
- Thin gold disks on a uniform background
- Alternative Forced Choice (AFC)
  - Observer has to guess the location of the disc
  - CDCOM software to score the presence/absence of disks
  - fit %_correct to get ‘threshold contrast’

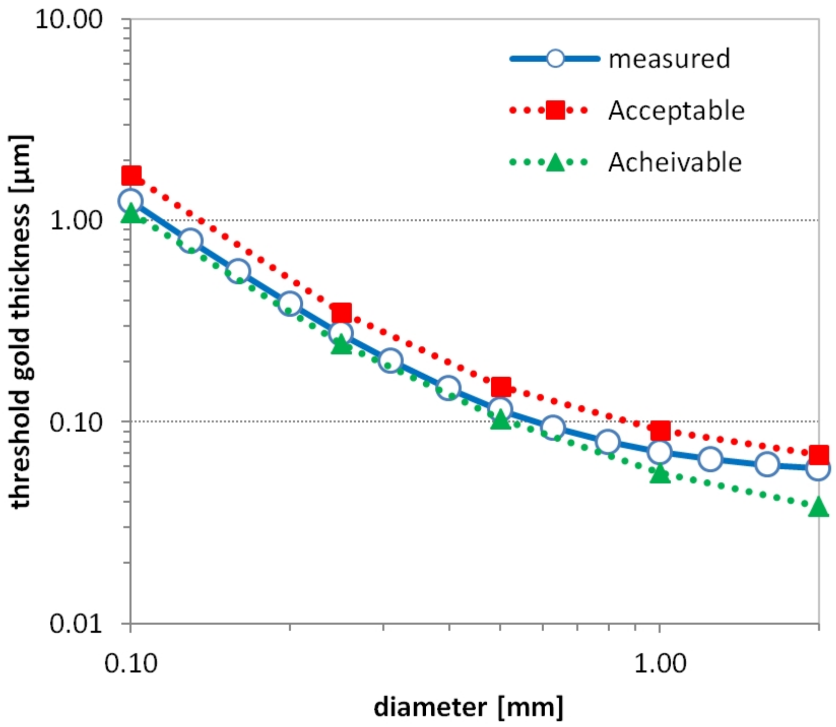




Bijkerk KR, Thijssen MAO, Arnoldussen ThJM, “Manual CDMAM-phantom type 3.4“ University Medical Centre Nijmegen

Test Object IQ - Threshold contrast-detail

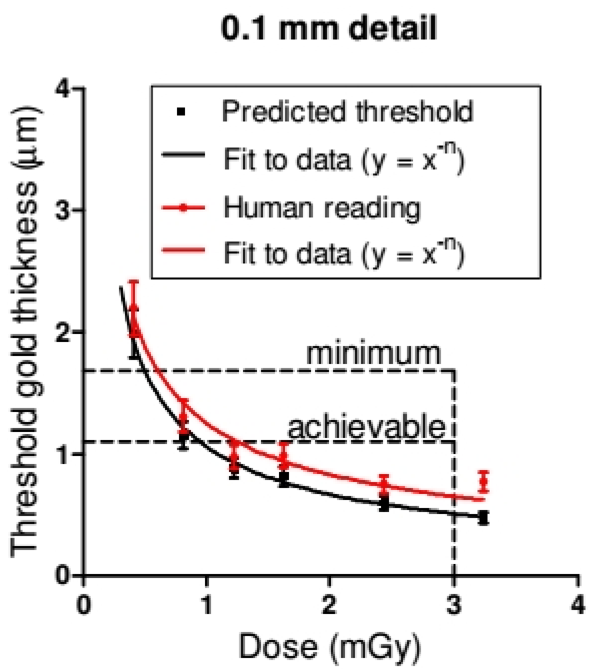
- CDMAM evaluation is ~system test
  - Test object at patient position
  - Imaged with scattering object
  - System selects spectrum (tube voltage, anode/filter combination)
- Excludes
  - Image processing, influence of anatomy on detection
- System has to pass the Acceptable level within the dose limit in order before it can be used for breast screening (EUREF)
  - Followed by many European countries



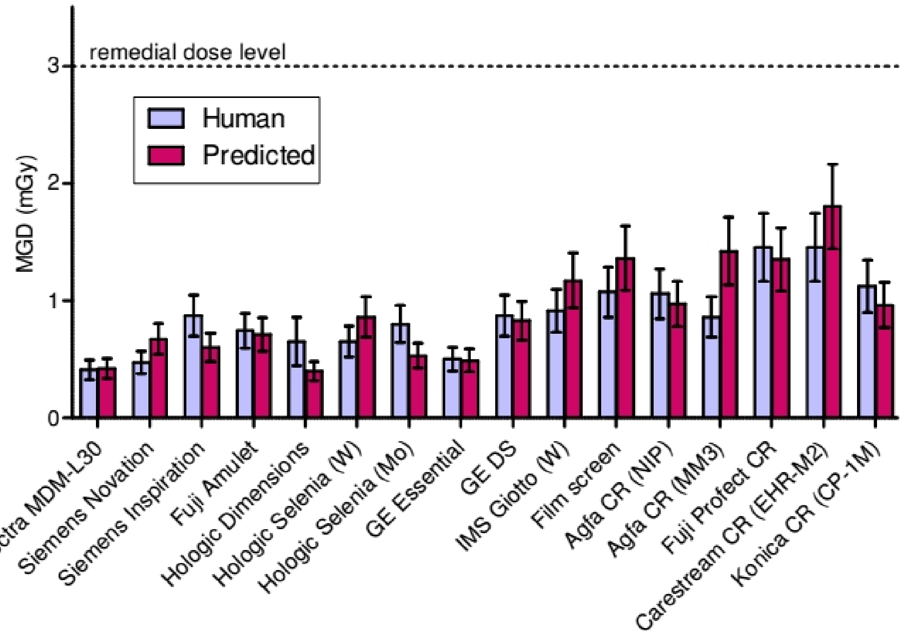
Young *et al* *SPIE* (2006)

Test Object IQ - Threshold contrast-detail

- Threshold gold thickness for 0.1 mm disk <= 1.68 µm (Acceptable) within the 3 mGy dose limit for 60 mm breast



Hologic Selenia

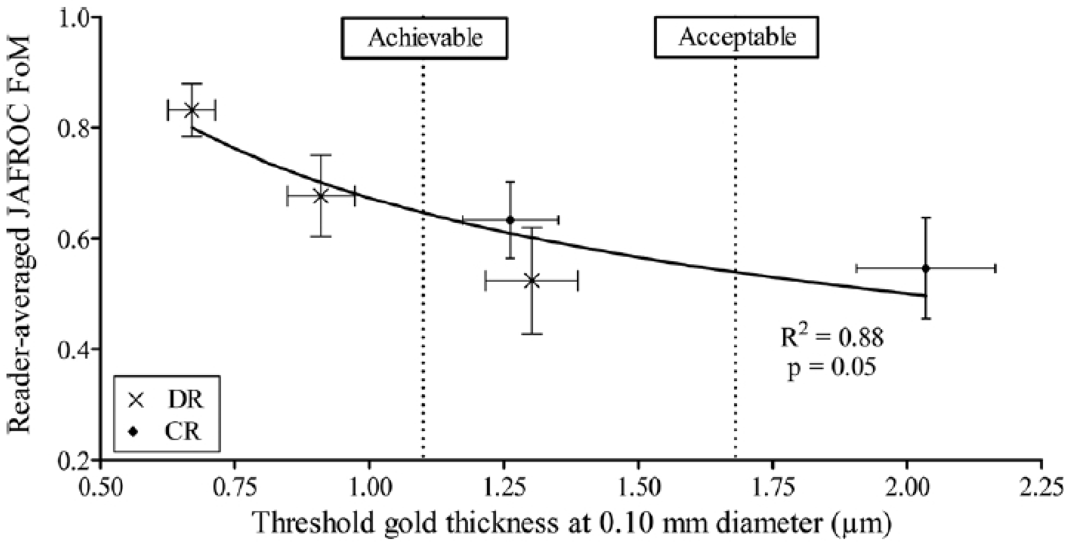


Compare systems

NHSBSP Equipment Report 1101 February 2011

Test Object IQ - Threshold contrast-detail

- But…the c-d test method [CDMAM] works
- Threshold gold thickness is related to clinical figure of merit for calcification detection
- Screening mammography has a relevant performance standard



Warren *et al* 2012 Effect of image quality on calcification detection in digital mammography *Med Phys*

Pros and Cons of the contrast-detail method

- Applicable to any system – don’t need access to special image types
- Curve is reasonably easy to interpret
- Assesses image SNR for digital (pixellated) imaging systems (not a line pair TO)

- Expensive test object
- Subjective: human observer scoring variability
- Learning effects (disc position)
- Test object variability (manufacturing tolerance)
- CDCOM - something of a “black box”
- Manual scoring can be time consuming (8 images…)

Derive Test Object performance from quantitative measurements ?

- Digital imaging systems
  - Access to pixel data via DICOM “For Processing” images
- We can calculate quantitative measures
  - Modulation Transfer Function (MTF)
  - Noise Power Spectrum (NPS)
  - Contrast (C)
- Combine these data in a human observer model
  - Choose a model that predicts human detectability results in low pass (typically) x-ray noise

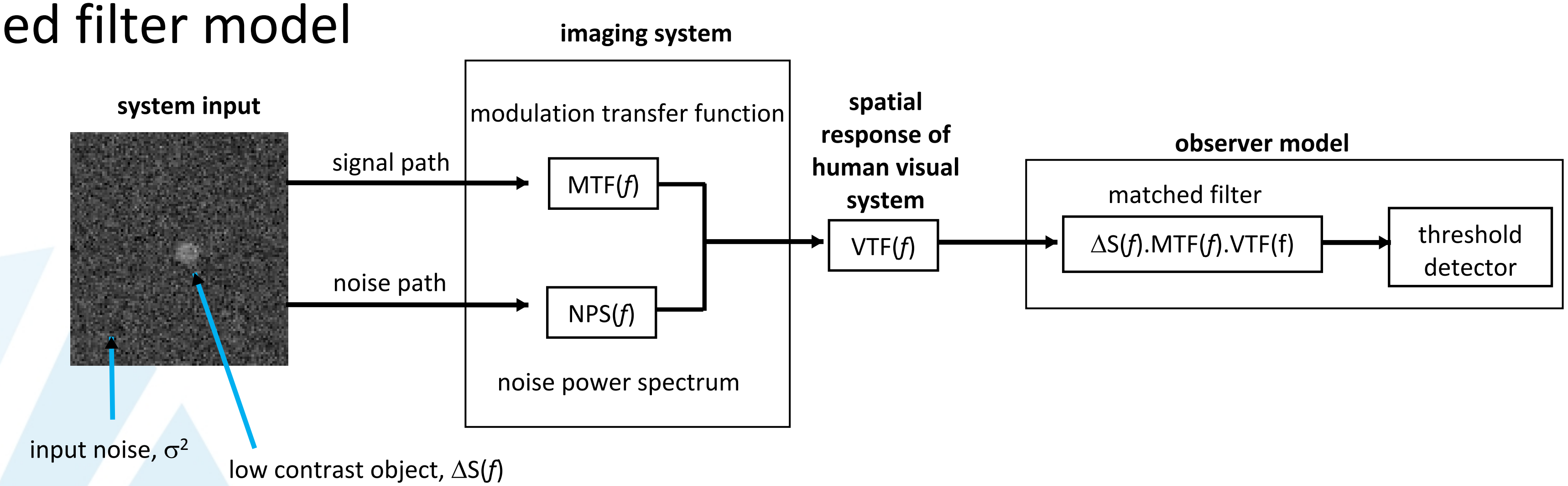
Aufrichtig R 1999 Comparison of low contrast detectability between a digital amorphous silicon and screen-film… Med. Phys. 26

What is a model observer?

- A metric, calculated from images generated by the imaging system
- Model observers can be formed from physical image quality parameters
  - in the spatial domain using matrix representation e.g. system covariance matrix
  - in the spatial frequency domain (with some strong assumptions) using Modulation Transfer Function (MTF), Noise Power Spectrum
- Ideal observer
  - Able to use all information available in carrying out the task
- Other observers
  - Cannot account for correlations in the noise, can be tuned to model human performance

Observer model

- Matched filter model



- The observer task is to detect a known (diameter, shape, position) object with signal spectrum $\Delta S(f)$, given the input noise variance σ^2
- The observer model is a template matched to the signal ('matched filter'), modified by the MTF(f) and the VTF(f)

L-N D Loo, K Doi and CE Metz 1984 A comparison of physical image quality indices and observer performance in the radiographic detection of nylon beads Phys. Med. Biol. 29 837-856

Observer model

- This is known as the non-prewhitened model observer with eye filter (NPWE)

$$d' = \frac{S}{N} = \frac{2\pi C \int_0^\infty \Delta S^2(f) MTF^2(f) VTF^2(f)\, f\, df}{\left(\int_0^\infty \Delta S^2(f) MTF^2(f) VTF^4(f)\, NNPS(f) f\, df\right)^{1/2}}$$

- the numerator (S) is the square root of the peak power resulting from a correct match
- the denominator (N) is the amplitude of the matched filter output for a given input noise σ^2, acted on by the system to give noise power NNPS(*f*)

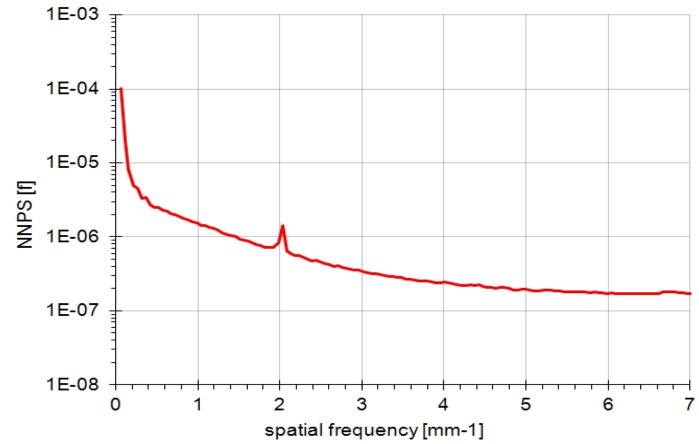
Observer model - components

Object signal (disc)

$$S(f) = \frac{d}{2}\frac{J_1(\pi d f)}{f}$$

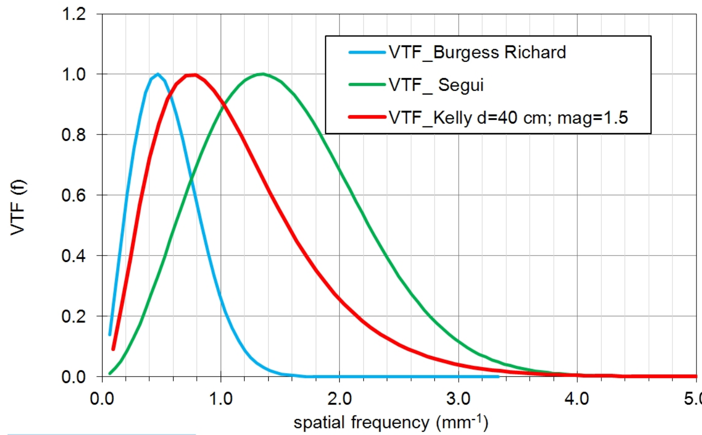
$$d' = \frac{2\pi C \int_0^\infty S^2(f) MTF^2(f) VTF^2(f)\, f\, df}{\left(\int_0^\infty S^2(f) MTF^2(f) VTF^4(f)\, NNPS(f)\, f\, df\right)^{1/2}}$$

NNPS



Visual Transfer Function VTF

$$VTF(f) = 29.5 f^2 \exp(-4f)$$

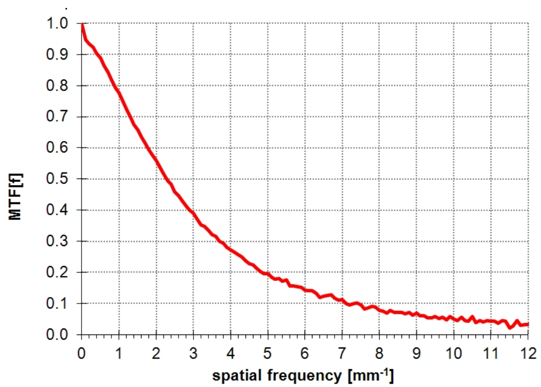


Contrast Al square

$$C = \frac{PV_{PMMA} - PV_{Al}}{PV_{PMMA}}$$

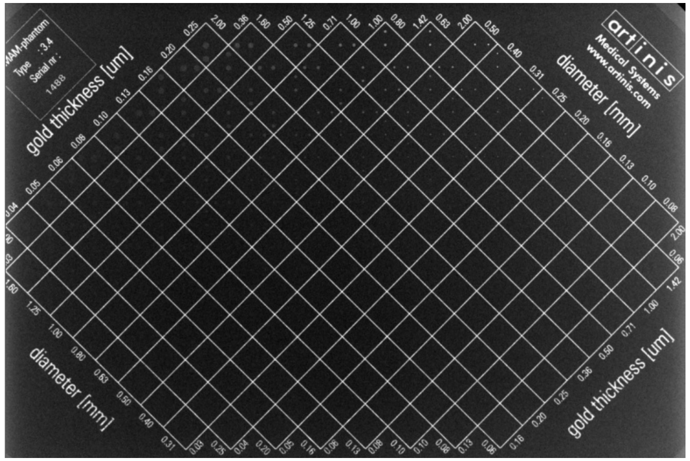
Presampled MTF

Dete
MTF

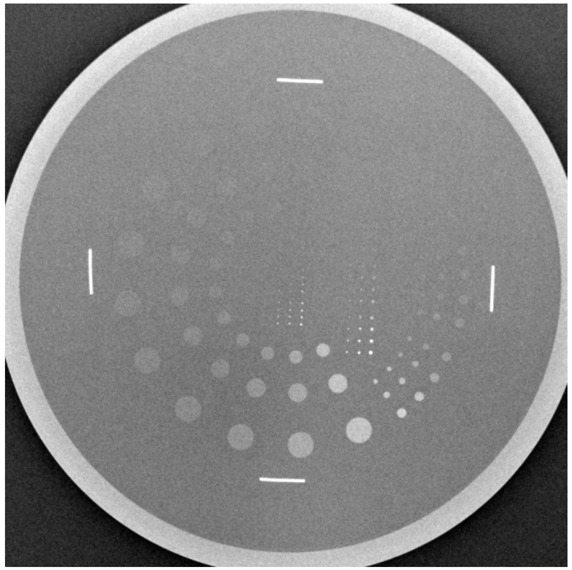


Kelly D H, Motion and vision II. Stabilized spatio-temporal threshold surface, J. Opt. Soc. Am. 69, 1340-1349, 1979

Model observer method for projection x-ray imaging



- Mammography imaging systems
  - Acquire contrast detail data as fn detector exposure with CDMAM c-d test object



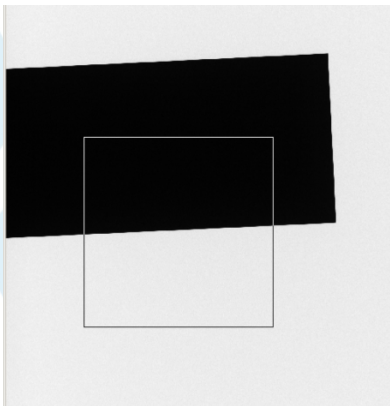
- General radiography imaging systems
  - Acquire contrast detail data as fn detector exposure with Leeds TO20 c-d test object
- For both modalities, measure quantitative image quality parameters
- Calculate d’ and compare with c-d result

Mammography

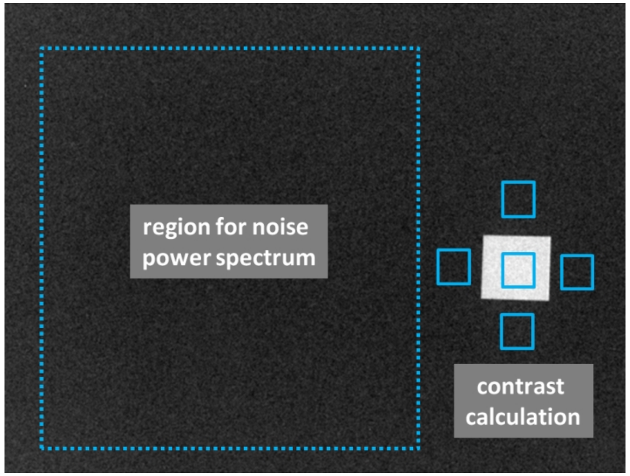
- Response function
  - 28 kV; 2 mm Al; 12 to 400 µGy
- Presampling MTF; edge at detector
- Contrast: 0,2 mm thick Al square
  - use AEC image (5 cm PMMA)
- NNPS from AEC image (5 cm PMMA)

Radiography

- Response function
  - 70 kV; 1 mm Cu; 0,25 µGy to 16 µGy
- Presampling MTF; edge at detector
- Contrast: 2 mm thick Al square
- NNPS from uniformly exposed image



MTF



**Contrast;
NNPS**

Calculate d’ – simple spreadsheet or program

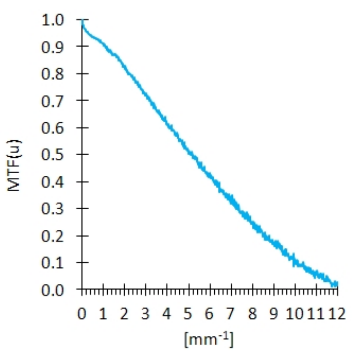
contrast

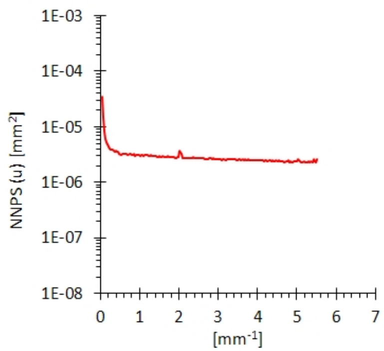
MTF

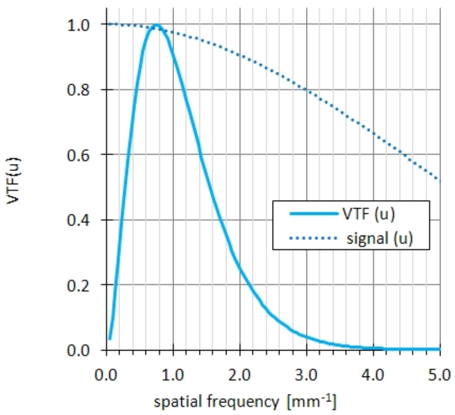
NNPS

measured Contrast: 0.144

NPWE d' for 0.1 mm diameter disc: *1.60*
NPWE d' for 0.25 mm diameter disc: *8.96*





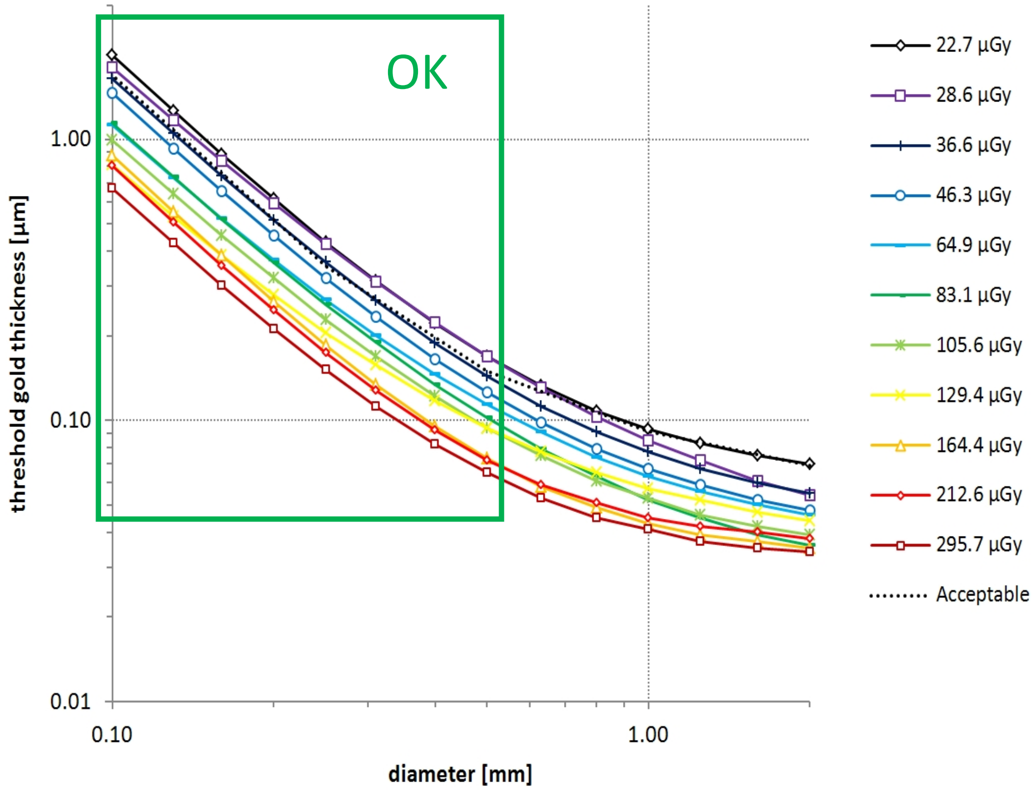


	full frequency	*interpolated*			*0.05*	*0.125*		*sum:*	*4.292E-05*
u (mm⁻¹)	*avg MTF(u)*	*avg MTF(u)*	*u (mm⁻¹)*	*NNPS*	*S² 0.1 mm(u)*	*S² 0.25 mm(u)*	*VTF (u)*	*du*	*NPWE num 0.1*
0.00	1.00	0.97	0.06	3.47E-05	6.17E-05	2.41E-03	0.03	0.044	1.71E-10
0.02	0.98	0.97	0.10	1.42E-05	6.17E-05	2.41E-03	0.10	0.040	2.26E-09
0.04	0.98	0.96	0.14	8.01E-06	6.17E-05	2.40E-03	0.18	0.048	1.19E-08
0.06	0.97	0.96	0.19	5.92E-06	6.16E-05	2.40E-03	0.28	0.049	4.03E-08
0.08	0.97	0.95	0.24	5.34E-06	6.16E-05	2.39E-03	0.39	0.044	8.83E-08
0.10	0.97	0.95	0.28	4.79E-06	6.16E-05	2.38E-03	0.49	0.044	1.64E-07
0.12	0.97	0.95	0.32	4.47E-06	6.15E-05	2.37E-03	0.58	0.042	2.54E-07
0.14	0.96	0.94	0.37	4.00E-06	6.15E-05	2.36E-03	0.66	0.049	4.27E-07
0.16	0.96	0.94	0.42	3.94E-06	6.14E-05	2.35E-03	0.75	0.049	6.15E-07
0.18	0.96	0.94	0.46	3.96E-06	6.14E-05	2.33E-03	0.82	0.044	7.39E-07
0.20	0.96	0.93	0.51	3.72E-06	6.13E-05	2.32E-03	0.87	0.045	9.29E-07
0.22	0.95	0.93	0.55	3.54E-06	6.12E-05	2.30E-03	0.92	0.046	1.13E-06
0.24	0.95	0.93	0.60	3.71E-06	6.11E-05	2.28E-03	0.95	0.047	1.34E-06
0.26	0.95	0.93	0.65	3.59E-06	6.11E-05	2.26E-03	0.98	0.044	1.42E-06
0.28	0.95	0.93	0.69	3.52E-06	6.10E-05	2.24E-03	0.99	0.047	1.65E-06
0.30	0.95	0.93	0.74	3.30E-06	6.09E-05	2.22E-03	1.00	0.049	1.89E-06
0.32	0.95	0.93	0.79	3.16E-06	6.08E-05	2.19E-03	1.00	0.047	1.92E-06
0.34	0.94	0.92	0.83	3.29E-06	6.06E-05	2.16E-03	0.99	0.045	1.86E-06
0.36	0.94	0.92	0.88	3.21E-06	6.05E-05	2.14E-03	0.97	0.044	1.85E-06
0.38	0.94	0.92	0.92	3.19E-06	6.04E-05	2.11E-03	0.95	0.045	1.91E-06
0.40	0.94	0.91	0.97	3.26E-06	6.03E-05	2.08E-03	0.93	0.047	1.98E-06
0.42	0.94	0.91	1.01	3.22E-06	6.01E-05	2.05E-03	0.90	0.046	1.88E-06
0.44	0.94	0.91	1.06	3.13E-06	6.00E-05	2.02E-03	0.87	0.045	1.78E-06
0.46	0.94	0.89	1.10	3.22E-06	5.99E-05	1.99E-03	0.84	0.047	1.77E-06
0.48	0.93	0.90	1.15	3.22E-06	5.97E-05	1.96E-03	0.81	0.047	1.70E-06
0.50	0.93	0.89	1.20	3.16E-06	5.95E-05	1.92E-03	0.77	0.045	1.51E-06
0.52	0.94	0.88	1.24	3.15E-06	5.94E-05	1.89E-03	0.74	0.045	1.41E-06
0.54	0.94	0.88	1.29	3.19E-06	5.92E-05	1.85E-03	0.70	0.046	1.34E-06
0.56	0.93	0.88	1.33	3.04E-06	5.90E-05	1.82E-03	0.67	0.046	1.25E-06
0.58	0.93	0.87	1.38	3.15E-06	5.88E-05	1.78E-03	0.63	0.046	1.14E-06
0.60	0.93	0.87	1.43	3.18E-06	5.87E-05	1.75E-03	0.59	0.045	1.00E-06
0.62	0.93	0.87	1.47	3.13E-06	5.85E-05	1.71E-03	0.56	0.045	9.15E-07
0.64	0.93	0.87	1.52	2.99E-06	5.83E-05	1.67E-03	0.53	0.047	8.78E-07
0.66	0.93	0.87	1.56	3.10E-06	5.81E-05	1.63E-03	0.50	0.047	7.91E-07
0.68	0.93	0.86	1.61	3.07E-06	5.78E-05	1.59E-03	0.46	0.047	6.96E-07

Mammography: validation against CDMAM

- 11 different digital mammography systems (CR, DR, photon counting)
- AEC technique (kV, A/F) for 5 cm PMMA
- Vary air kerma at detector (K) around the AEC operating air kerma point
- Measure detectability (d') and threshold gold thickness (T)
- At each K setting take 8 CDMAM images
- Score using Erica[2] software (LUCMFR, Leuven)

$$T^2\, d^2\, K \approx constant$$

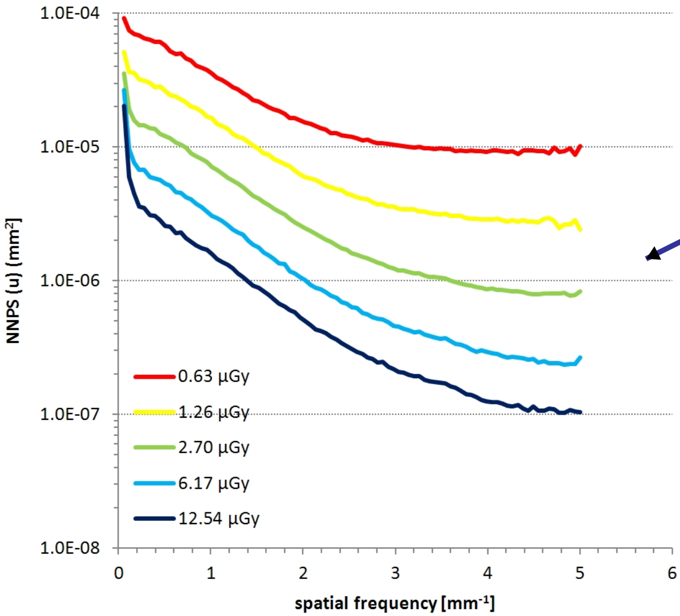


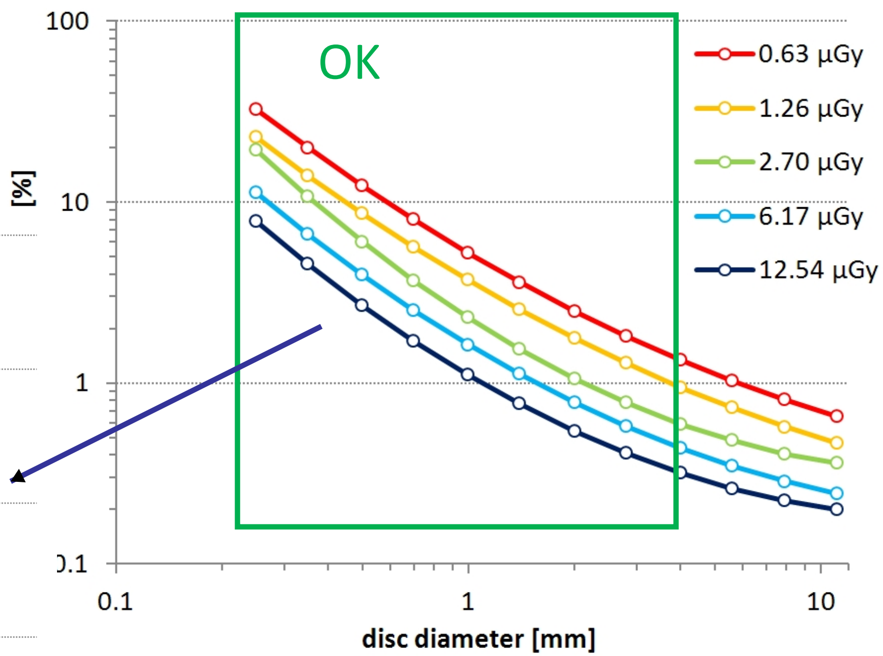
Example for GE Essential

Radiography: validation against TO20

- 4 different digital systems (CR, DR)
- Manual technique (70 kV, W/Al + 1mm Cu)
- Vary air kerma at detector (K) around standard operatin
- Measure detectab contrast (C%)
- At each K setting a images; human sc

$$T^2 d^2 K \approx constant$$

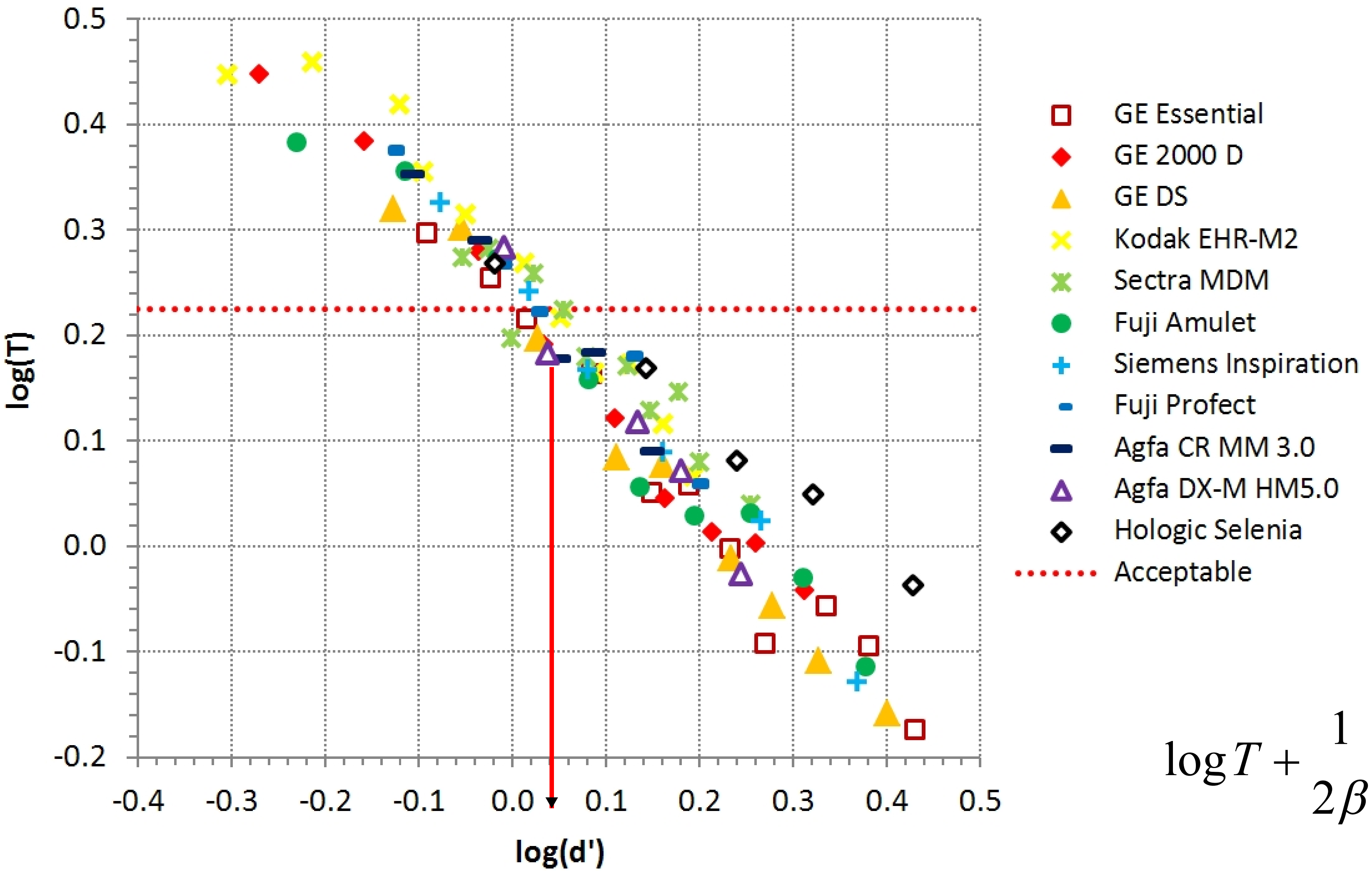




Example for Carestream DRX 1C

falling NNPS ~ lower threshold contrast

Mammography: threshold gold thickness vs d’ [0,1 mm disc]



$$\log T + \frac{1}{2\beta}\log d' = constant$$

Monnin et al Image quality assessment in digital mammography: part II. NPWE as a validated alternative for contrast detail analysis
Phys. Med. Biol. 56, 4221- 4238 (2011)

Mammography: threshold d’ index for 0.1 mm disc

- Perform linear fit to these log-log data and establish threshold d’ for a given threshold gold thickness

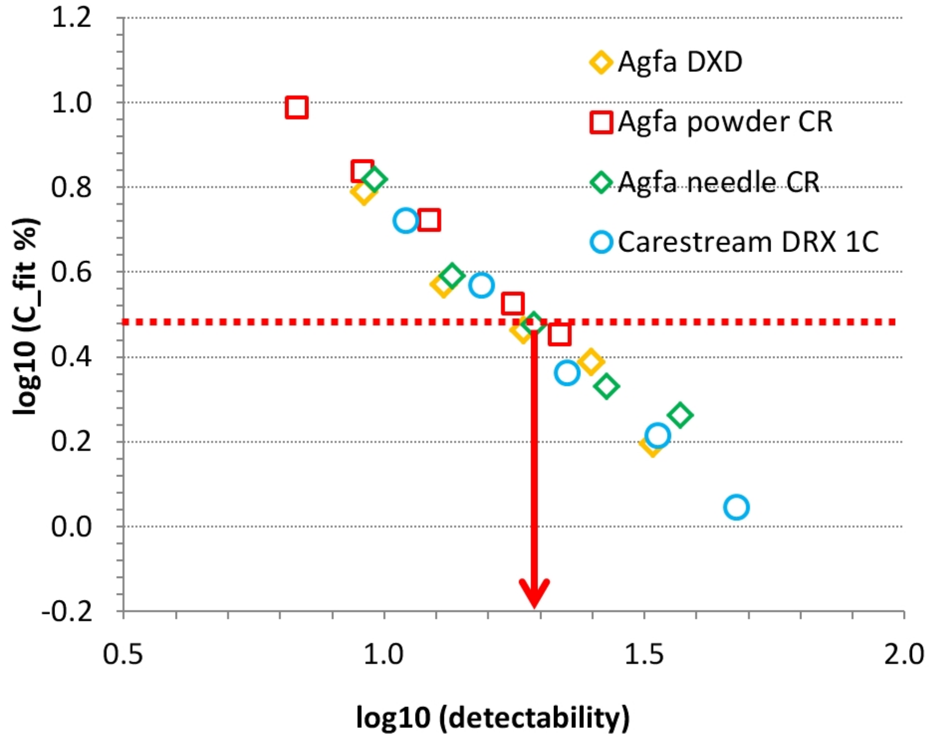
Imaging system	Tube potential anode/filter	Image contrast	Fitted slope (σ)	correlation coefficient (R^2)	Threshold *d’*	Difference from target *d’* = 1.05 [%]	Threshold detector air kerma [μGy]
Agfa MM 3.0R	29 kV - Mo/Rh	0.125	-0.993 (0.099)	-0.99	**1.06**	**0.5**	100
Agfa HM 5.0	29 kV - Mo/Rh	0.131	-1.111 (0.116)	-0.98	**1.06**	**1.2**	65
Carestream EHR-M3	28 kV - Mo/Mo	0.130	-0.906 (0.061)	-0.99	**1.05**	**-0.1**	83
Carestream SNP-M1	28 kV - Mo/Mo	0.140	-0.895 (0.094)	-0.97	**1.09**	**3.5**	52
Fuji Amulet	29 kV - W/Rh	0.124	-0.862 (0.061)	-0.98	**0.99**	**-5.5**	41
Fuji Profect	27 kV - Mo/Rh	0.139	-0.900 (0.100)	-0.98	**1.08**	**3.0**	66
GE Senographe 2000 D	28 kV - Rh/Rh	0.133	-0.903 (0.037)	-0.99	**1.00**	**-5.1**	58
GE Senographe DS	29 kV - Rh/Rh	0.133	-0.970 (0.039)	-0.99	**0.98**	**-6.4**	46
GE Senographe Essential	29 kV - Rh/Rh	0.139	-0.911 (0.056)	-0.98	**0.99**	**-6.2**	32
Hologic Selenia Dimensions	29 kV - W/Rh	0.129	-0.906 (0.061)	-0.99	**1.08**	**3.1**	42
Sectra MDM	35 kV - W/Al	0.085	-0.743 (0.080)	-0.95	**1.07**	**1.4**	41
Siemens Inspiration	28 kV - W/Rh	0.140	-0.988 (0.053)	-0.99	**1.07**	**1.5**	45

Expect fitted slope to be 1.0 for quantum limited systems as $(2\beta)^{-1} \sim 1.0$

For 0.1mm disc T = 1.68 μm □ d’ = 1.05

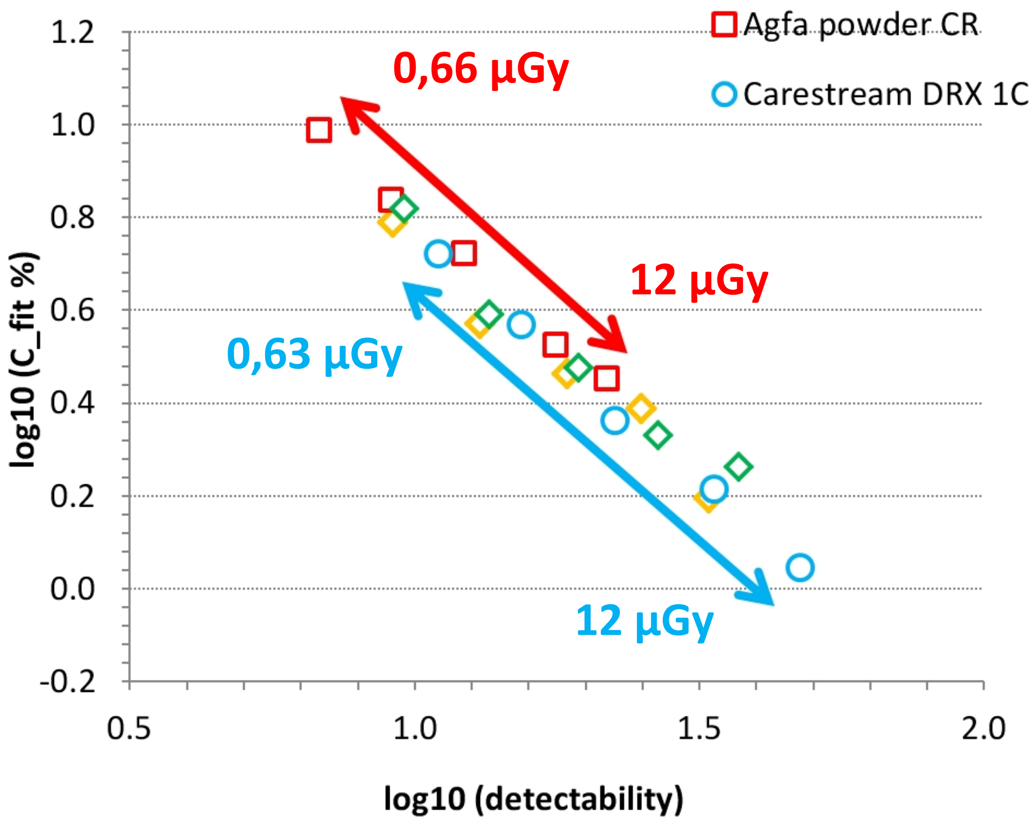
Radiography: C% vs d' for 0,5 mm disc

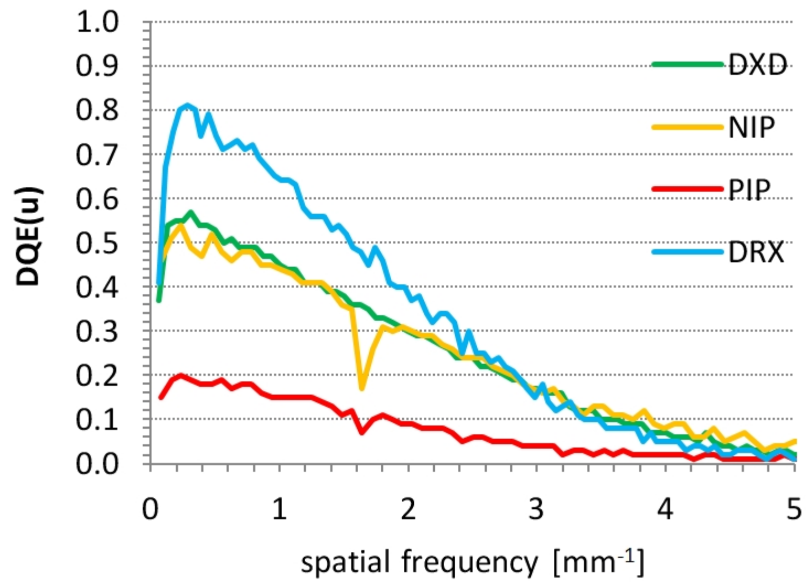
- Closely follows the result for mammography systems
- Fitted slope is 1,07 (expect 1.0 for quantum limited systems)
- Not mammography – DR detectors are used for a wide range of imaging tasks
- Some detectors are better than others….
    - Higher IQ for same dose
    - Same IQ for lower dose



Radiography: C% vs d' for 0,5 mm disc

- Large differences in C% and detectability are a reflection of differences in detector efficiency



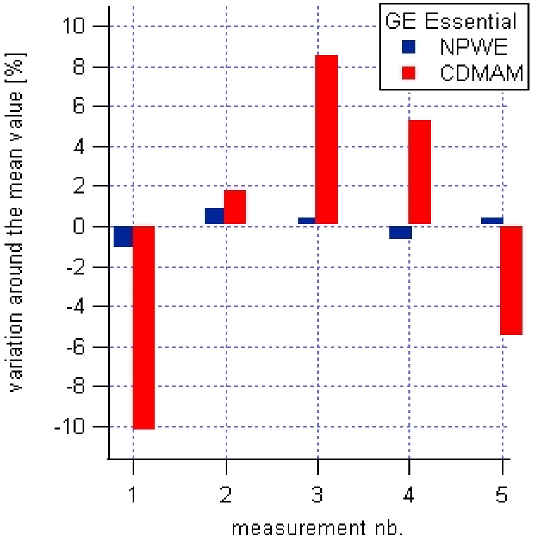


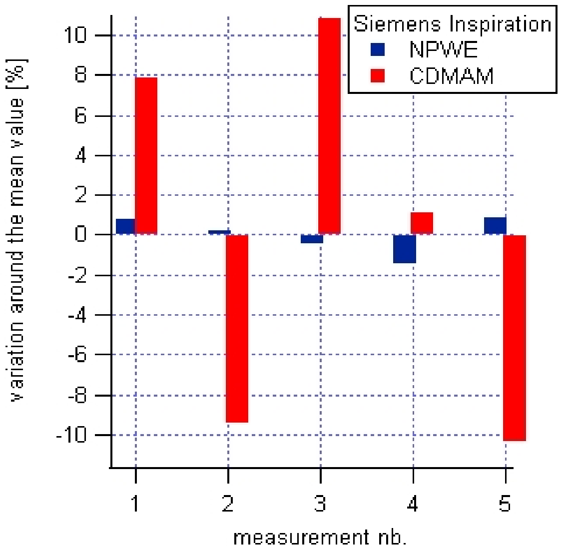
Detectability index d’: explicit dependencies

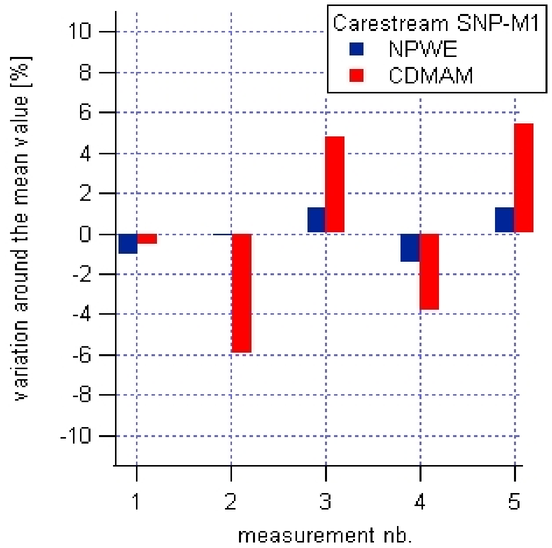
- In the c-d method, all aspects of performance are wrapped up in a single threshold value for some disc diameter
  - It may not be clear what aspect is limiting performance
- The d’ method uses quantitative measurements in a system or detector context
  - Contrast (beam energy and scatter rejection)
  - Sharpness (presampling MTF)
  - Noise (NPS) – detector DQE, exposure at detector
- We hope this gives better insight into why systems perform well or poorly

Robustness and repeatability

- Model Observer method = better reproducibility
  - variability of d’ ⍰ 2% against T ⍰ 10% for the CDMAM 0,1 mm disc
  - No partial volume effects
  - Larger ROI for contrast and noise evaluation







Summary Points

- True task based evaluation of DR imaging systems is not done routinely by physicists
  - Time consuming (physicists and radiologists)
  - Suited mainly to detection and some characterization tasks
  - If required then we have to look for alternatives
- Quantitative methods are ideally suited to QC
  - checking detector performance, tracking changes
- These can be used to construct measures of detectability that correlate with test object evaluations using simple tasks

Summary Points

- The challenge is to find ways of making task based IQ evaluation ‘in the room’
  - Eg. troubleshoot cases where the clinicians say that system IQ is inadequate for their tasks…
- OPTIMIZATION of system for given tasks
  - Is the level of sharpness, noise (dose…), contrast, recursive filtering etc set at correct level for tasks performed on the system?
  - Difficult to do this using e.g. c-d test objects

QC - Repeatability of quantitative measures

- Measure all the parameters required to calculate DQE on 5 consecutive days for an FFDM system
- Hologic Selenia
- 28 kV, W/Rh, 2 mm Al at the tube, grid out
- 1 image for MTF, 1 image for NNPS

28 kV W/Rh	Mon	Tues	Wed	Thurs	Fri	cov (%)
air kerma output factor :	3.00	2.96	2.99	2.99	2.99	*0.5*
response fn offset :	50.86	49.69	50.19	49.46	49.39	*1.2*
response fn gradient :	2.74	2.78	2.77	2.79	2.78	*0.7*
true air kerma for 100 µGy target :	97.5	96.2	97.2	97.2	97.2	0.5

cov = coefficient of variation (=σ/mean)